1-1-2014

Cognitive Enhancement With Stimulants: Effects and Correlates

Irena Ilieva
University of Pennsylvania, irena.p.ilieva@gmail.com

Cognitive Enhancement With Stimulants: Effects and Correlates

Abstract
Does non-medical use of prescription stimulants improve healthy cognition? What distinguishes healthy users of ADHD medication from their peers? The present project examined stimulants' cognitive enhancement effects and the psychological profile of non-medical stimulant users. Study 1, a double-blind, placebo-controlled experiment, found no enhancing effect of mixed amphetamine salts (Adderall) on healthy participants' inhibitory control, working memory, episodic memory, convergent creativity, perceptual intelligence, and a standardized achievement test. No moderating effects of baseline performance or COMT genotype were detected. Despite the lack of enhancement observed for most measures and most participants, participants nevertheless believed their performance was more enhanced by the active capsule than by placebo. In Study 2, we conducted a meta-analysis to determine whether stimulants' cognitive enhancement potential is truly non-existent or simply small. Based on 47 double-blind, placebo-controlled experiments, we found evidence for small effects of amphetamine and methylphenidate on inhibitory control, working memory and episodic memory. Given the absence of conclusive evidence for practically significant stimulant effects in healthy people, we conducted Study 3 to infer about the motives for use from users' psychological profile. Non-medical stimulant use appeared more strongly related to individuals' perceived attention functioning than to their objectively measured attentional performance. Users reported lower motivation during the laboratory attention test and described their everyday study habits as poorer than a control group with no history of stimulant use. Taken together, these data imply that enhancement users struggle with below-average functioning in one or several cognitive, affective and behavioral domains, possibly seeking stimulants to compensate for these problems.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Psychology

First Advisor
Martha J. Farah

Keywords
executive functions, neuroethics, unprescribed stimulants, working memory

Subject Categories
Psychology

COGNITIVE ENHANCEMENT WITH STIMULANTS: EFFECTS AND CORRELATES

Irena Parvanova Ilieva

A DISSERTATION

in

Psychology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

_____

Martha Farah

Professor of Psychology

Graduate Group Chairperson

_____

John Trueswell

Professor of Psychology

Dissertation Committee

Geoffrey Goodwin (Chair), Associate Professor of Psychology

Robert DeRubeis, Professor of Psychology

COGNITIVE ENHANCEMENT WITH STIMULANTS: EFFECTS AND CORRELATES

ACKNOWLEDGEMENT

I am profoundly grateful to my adviser Martha Farah for the opportunity to train under her exceptional mentorship. Her wise guidance, thought-provoking questions, constructive feedback and unending supportiveness have given me a perfect milieu for professional growth. I am incredibly thankful to my committee members, Geoff Goodwin and Rob DeRubeis for their expertise and support. Their insightful feedback has been truly invaluable to my research training. I am beyond fortunate to have had the opportunity to learn from Martha and my committee, who, aside from being remarkable role models, have been highly invested in fostering my development.

I am grateful to Melissa Hunt, Sharon Thompson-Shill and Anjan Chatterjee for stimulating discussions on some aspects of this project. I wish to thank my lab-mates and fellow students – Gwen Lawson, Daniel Hackman, Teresa Pegors, Marieta Pehlivanova, Pavel Atanasov, Marcelo Mattar, Eli Tsukayama, Cayce Hook, Omar Butt and Kelly Matula – for their excellent suggestions about parts of this work. Thanks to my lab-mates Cayce Hook, Joe Boland, Melissa Johnston and Connor McLaren for their enormous help with data collection and data management.

ABSTRACT

COGNITIVE ENHANCEMENT WITH STIMULANTS:

EFFECTS AND CORRELATES

Irena P. Ilieva

Martha J. Farah

Does non-medical use of prescription stimulants improve healthy cognition? What distinguishes healthy users of ADHD medication from their peers? The present project examined stimulants' cognitive enhancement effects and the psychological profile of non-medical stimulant users. Study 1, a double-blind, placebo-controlled experiment, found no enhancing effect of mixed amphetamine salts (Adderall) on healthy participants' inhibitory control, working memory, episodic memory, convergent creativity, perceptual intelligence, and a standardized achievement test. No moderating effects of baseline performance or COMT genotype were detected. Despite the lack of enhancement observed for most measures and most participants, participants nevertheless believed their performance was more enhanced by the active capsule than by placebo. In Study 2, we conducted a meta-analysis to determine whether stimulants' cognitive enhancement potential is truly non-existent or simply small. Based on 47 double-blind, placebo-controlled experiments, we found evidence for small effects of amphetamine and methylphenidate on inhibitory control, working memory and episodic memory. Given the absence of conclusive evidence for practically significant stimulant

effects in healthy people, we conducted Study 3 to infer about the motives for use from users' psychological profile. Non-medical stimulant use appeared more strongly related to individuals' perceived attention functioning than to their objectively measured attentional performance. Users reported lower motivation during the laboratory attention test and described their everyday study habits as poorer than a control group with no history of stimulant use. Taken together, these data imply that enhancement users struggle with below-average functioning in one or several cognitive, affective and behavioral domains, possibly seeking stimulants to compensate for these problems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

PREFACE

Cognitive enhancement refers to the use of psychiatric medication to optimize *healthy* cognition. Pharmacological approaches to the enhancement of cognition have sparked interest for decades (Rassmussen, 2008). Currently, a popular enhancement approach involves the use of ADHD medications, most commonly Adderall and Ritalin, by healthy young people *without* ADHD who seek to compete successfully in school or at work. The present work discusses the effects, correlates and implications of this practice.

Relative to other risky behaviors, unprescribed stimulant use is highly prevalent. Among college students, the prevalence of lifetime use, though as low as 2 % on some campuses (Nowak et al., 2007), can exceed 35% (Low & Gendazcek, 2002) on others. Most popular among college students, stimulants are also used non-medically by surgeons, lawyers, financial traders and professional academinics (Franke et al., 2014, Talbot, 2009, Sahakian & Morein-Zamir, 2011), especially during times of high pressure for productivity (DeSantis et al., 2008). Aside from the United States, use has been documented in Canada, Western Europe and Australia (Sahakian and Morein-Zamir, 2007; Franke, 2011; Castaldi et al., 2012; Partridge et al., 2013).

Not only is interest in the non-medical use of drugs like Adderall and Ritalin high, but access to these medications is relatively easy. When diagnosing clinicians rely only on self-report, symptoms of ADHD can be easily faked to obtain a prescription. Stimulants may be prescribed even without a diagnosis if the physician views medication

as a tool for alleviating distress independent of the presence of a diagnosable condition (Greely et al., 2008). Between 14.7 % and 26% of people with ADHD have indicated diverting their medication to undiagnosed peers (e.g., Poulin et al., 2007; Poulin et al., 2001). ADHD medications are affordable, especially if used only occasionally (DeSantis et al., 2008). The availability and wide interest in unprescribed stimulants underscore the importance of examining their enhancement potential.

Knowledge about stimulants' enhancement effects can benefit both individual users and society at large. In theory, cognitive enhancement holds substantial promise for progress in virtually all areas of life. To name a few examples, enhancement tools can improve medical professionals' decision-making, aircraft pilots' vigilance, students' grasp of the material. Incrementally, cognitive enhancement may even facilitate scientific and medical progress, for instance, by leading to the discovery of cures for previously intractable diseases (Roache, 2008). An assessment of stimulant-related benefits is a crucial first step towards determining if these hypothetical perspectives are realistic.

Additionally, data on the advantageous effects of cognitive enhancers can help scientists, ethicists and policy-makers weigh the benefits against the costs. Several risks of non-medical stimulant use require consideration, including the potential for abuse, dependence and cardiovascular problems (Chatterjee, 2009), the possibility of detrimental stimulant effects on some facets of cognition (e.g., de Wit et al., 2014), the uncertainty about stimulants' underresearched long-term side effects (Farah et al., 2004), the potential for coercion into use and the danger of uneven access for different social classes (Farah et al., 2004). Assessing the actual potential of currently available enhancement agents is an important step towards a cost-benefit analysis.

One source of knowledge about stimulants' enhancement effects is non-medical users' own reports. According to past research, these reports depict ADHD medications as beneficial for healthy attention, memory, intelligence, alertness and energy (Rabiner et al., 2009, DeSantis et al., 2008). These intuitions may be riddled with optimistic misconceptions. Evidence is accumulating that the general public may harbor an overly rosy view of cognitive enhancers' promises. For instance, some users inappropriately discount the risks of self-medication, based on the false assumption that FDA approval alone guarantees a drug's safety even without medical supervision (DeSantis et al., 2010). College students overestimate the actual prevalence of use on campuses (McCabe et al., 2008), potentially leading to a perception of use as more normative, and possibly, as more beneficial. Media portrayals of enhancement use may be facilitating an uncritical belief in ADHD medications' positive effects: an analysis of media coverage on the topic showed that 85% of retrieved texts included no reference to the scientific literature on enhancement effectiveness. These texts, while almost invariably emphasizing advantages of enhancement, mentioned the dangers only occasionally (Partridge et al., 2011). Taken together, these factors raise the possibility that users' perceptions of enhancement stimulants' effects may be positively biased.

Laboratory experiments on stimulants' effects on objective cognitive tests can provide more conclusive estimates of these medications' promise. Yet, despite decades of relevant research, this goal has remained elusive. As will be discussed in Studies 1 and 2, double-blind, placebo-controlled research on these medications' effects on various facets of cognition (e.g., episodic and working memory, inhibitory control, other executive functions) has yielded a mixture of positive and null results. The explanation

for these inconsistencies is unclear: research limitations might be masking practically significant effects; stimulant effects might be overall small; or the literature might be reflecting a combination of true null effects and selectively published false positive findings. Moreover, as elaborated in Study 1, assessment of the individual differences moderators of stimulant enhancement effects have been inconclusive because conducted in small samples with designs allowing alternative interpretations. Thus, the effects of prescription stimulants on healthy cognition have remained an open question.

An examination of stimulants' costs and benefits also has to take into account the psychological profile of users. Cognitive and affective problems experienced by users at baseline may impact stimulants' effects. For instance, it has previously been suspected that users suffer from untreated ADHD: such a pattern would increase the probability of stimulant benefits, but also increase the risks. Self-medication could demotivate appropriate help seeking for a potentially impairing condition. True pathology might lead to more frequent unpresecribed medication use, in turn leading to more exposure to medication risks without access to medical supervision. Thus, an assessment of the psychological characteristics of users is important component of evaluating the costs and benefits of cognitive enhancement. As will be discussed in Study 3, conclusive data on important aspects of users' profile is still unavailable.

Study 1 of the present dissertation assesses the objectively measured and subjectively perceived effects of mixed amphetamine salts (Adderall) on healthy memory, inhibitory control, intelligence, creativity and standardized test performance in a double-blind, placebo-controlled trial. The experiment also evaluates the role of two of these effects' candidate moderators: baseline cognitive performance and COMT

genotype. Study 2 is a meta-analysis of amphetamine and methylphenidate's effects on inhibitory control, working memory and episodic memory, which incorporates an assessment of publication bias. Study 3 examines real-world enhancement users' psychological profile, comparing their attention functioning (self-reported and objectively measured), study habits and motivation to a control group who has never used ADHD medication.

CHAPTER 1

OBJECTIVE AND SUBJECTIVE COGNITIVE ENHANCING EFFECTS OF MIXED AMPHETAMINE SALTS IN HEALTHY PEOPLE

Cognitive enhancement refers to the use of neuropsychological drugs, most commonly psychostimulants such as amphetamine and methyphenidate, by cognitively normal, healthy people to improve cognitive function. Evidence suggests that enhancement is a common practice and may be gaining in popularity. A study on a large 2001 sample of undergraduate programs including institutions of different size, location, religious affiliation and private/public status, showed an almost 7% lifetime prevalence of nonmedical stimulant use (McCabe, Knight, Teter & Wechsler, 2005). Although this study did not distinguish between cognitive enhancement and other nonmedical uses, more recent surveys of college students have done so and indicate that cognitive enhancement is the primary motivation for most students using stimulants (e.g., DeSantis, Webb, & Noar, 2008; see Smith & Farah, 2011, for a review). These more recent studies also indicate substantially larger proportions of students using prescription stimulants compared to the McCabe and colleagues' estimates, although the samples have been smaller and less representative. Aside from college students, enhancement use of stimulants has also been reported among professionals from various fields (e.g., lawyers, journalists, Madrigal, 2008; Maher, 2008; Talbot, 2009).

Stimulants' Actual Cognitive Enhancement Effects

One possible reason for the growing enhancement use of stimulants is that the drugs truly improve cognitive abilities such as learning and executive function, presumably through their effects on catecholamine neurotransmission (Meyer & Quenzer, 2005). Yet, in the aggregate, the evidence supporting stimulants' beneficial effects on healthy cognition is mixed. For example, Chamberlain, Robbins, Winder-

Rhodes, Muller, Sahakian, Blackwell and Barnett (2010) reviewed studies in which CANTAB tasks had been used to assess stimulant effects in patients and healthy control participants. They concluded that "acute doses of medication improved aspects of cognition, though findings were more consistent in subjects with ADHD than in healthy volunteers." Reviewing the literature on the cognitive effects of methylphenidate, Repantis, Schlattmann, Laisney & Heuser (2010) state that they were "not able to provide sufficient evidence of positive effects in healthy individuals from objective tests." Similarly, Hall and Lucke (2010) state that "There is very weak evidence that putatively neuroenhancing pharmaceuticals in fact enhance cognitive function." An even stronger view was presented by Advokat (2010), whose reading of the literature led her to suggest that "studies in non-ADHD adults suggest that stimulants may actually impair performance on tasks that require adaptation, flexibility and planning."

Most recently, Smith and Farah (2011) surveyed more than fifty experiments on the effects of amphetamine and methylphenidate on a wide array of cognitive functions, including memory (episodic memory, procedural memory and probabilistic learning) and executive functions (working memory, cognitive control) in healthy young adults. They discovered a roughly even mixture of significant enhancement effects and null findings overall. Studies on episodic memory tended to show an enhancing effect of stimulants when retention intervals were longer than an hour, whereas evidence for enhancement of other functions was less clear. For executive functions (including inhibitory control, working memory and other executive functions) many studies reported significant enhancing effects but some did not. In addition, when found, these effects were sometimes qualified by complex interactions between the order of drug and placebo administration, participants' cognitive performance on placebo, and participants'

genotypes. The possibility that other null results have been found but not published (publication bias, also known as the "file drawer effect") must be considered. In sum, a number of recent reviews have concluded that the cognitive enhancement potential of stimulants has not received firm empirical support.

Several factors may explain the inconsistency between users' beliefs that stimulants enhance cognition and the equivocal evidence for these effects. One possibility is that the assessment of enhancement effects in the laboratory has been impeded by problems such as unmeasured moderators, poor measurement of moderators or low statistical power. These would be especially serious challenges to research in this area if the effects of stimulants are small and dependent on individual differences. Another possibility is that stimulants create a subjective perception of enhancement, possibly more salient and wide-spread than the actual effects. The rest of this section will elaborate on these potential explanations.

Challenges in Assessing the Enhancing Effects of Stimulants

Among the challenges standing in the way of settling the question of stimulants' enhancement potential are the following four. The majority of published studies fail to meet any of these challenges, and no study has so far been designed to address all four. These challenges motivate the design of the present double-blind, placebo controlled, cross-over trial on the cognitive enhancement effects of mixed ampheramine salts (MAS, brand name Adderall).

Moderation of enhancement effects by individual differences. One reason why previous research may have failed to detect significant evidence for enhancement is that stimulants may be effective for some individuals but not for others. Thus, studies that have not measured or analyzed the effect of moderating individual differences may have erroneously concluded that the effects are small or nonexistent. One candidate

moderator is individuals' endogenous dopamine activity. The relationship between dopamine activity and cognitive performance is believed to follow an inverted U-shaped curve, in which intermediate dopamine levels are optimal for cognitive performance, whereas low and high levels are detrimental (Robbins & Arnsten, 2009). Therefore, individuals at different starting points on this curve would benefit differentially from the increase of dopamine activity caused by a dose of stimulant. Individuals with sub-optimal baseline dopamine levels would be moved upward on the curve to higher cognitive performance. By contrast, individuals with high baseline dopamine, standing at the peak or on the downward-sloping portion of the curve, would move downward in cognitive performance.

Several studies have provided evidence for the moderation of stimulant effects by endogenous dopamine activity, as indexed by participants' Catechol *O*-methyltransferase (COMT) genotype. A common polymorphism of the COMT gene determines the activity of the COMT enzyme, which breaks down dopamine and norepinephrine. Hence, the COMT genotype influences the level of synaptic dopamine. Mattay and colleagues (2003) have shown that individuals whose COMT genotype is associated with higher endogenous dopamine show less enhancement by amphetamine and in certain tasks may actually perform worse on the drug.

Another possible moderator of amphetamine's cognitive enhancing effects is cognitive ability. Several studies have found that participants who perform worse than average when on placebo are more likely to be enhanced by stimulants (Farah, Haimm, Sankoorikal & Chatterjee, 2008; De Wit, Crean & Richards, 2000; de Wit, Enggasser & Richards, 2002, Mattay et al., 2000; Metha, Owen, Sahakian, Mavaddat, Pickard & Robbins, 2000). Findings of both COMT-moderated and performance-moderated enhancement suggest that some of the null results in literature may result from a

mixture of true enhancing effects for some individuals and absent or even reversed effects for others. Measurement of these two potential moderating factors is therefore crucial for determining the true enhancement potential of stimulant drugs. In the present study we measure both.

Regression to mean and measurement of baseline performance. Baseline performance, as a moderator of enhancement, has typically been indexed by performance on placebo. This measure is problematic because of the phenomenon of regression to the mean. To the extent that there is measurement error in the data, participants who score well in the placebo condition would be expected to score less well on average in a different session, and participants who score poorly in the placebo condition would be expected to score somewhat better on average in a different session. Consequently, even in the absence of moderation by baseline, placebo scores may appear to moderate the difference between drug and placebo purely due to regression to the mean. For this reason, we obtain a measure of baseline ability that is independent from participants' performance on drug and placebo.

Moderation by order of drug administration. Some previous within-subjects trials on the effects of stimulants on cognition have unexpectedly revealed a third moderator of enhancement effects. In particular, significant enhancement effects on three different tasks have been observed when the drug was administered before placebo, but not after (Elliott, Sahakian, Matthews, Bannerjea, Rimmer & Robbins, 1997). Such moderation is difficult to interpret; it might reflect a specificity of stimulant effects to novel tasks, or a specificity to more difficult tasks, or it may be a type II error. If order is not controlled and analyzed in within-subjects studies, the effects of stimulants could be inflated or diluted. Between-subjects studies are not free of this problem, as all participants effectively receive the drug or placebo first. If stimulant

effects are fleeting, then single-session between-subjects studies would overestimate the effectiveness of the drug. Accordingly, in the present study we control for the order of drug administration both experimentally (i.e., by counterbalancing the variable between participants) and statistically.

Statistical power. Insufficient statistical power to detect practically significant effects has been a major obstacle to discovering stimulants' cognitive enhancing properties. Most of the experiments reviewed by Smith & Farah (2011) used samples of fewer than 40 participants, many with between-subjects designs. The present within-subject study's sample size of 46 was chosen to give us 95% power to detect a medium-size effect (Cohen's d = 0.5) on any single measure.

Perceptions of Enhancement

Another way to explain the discrepancy between the rising enhancement use and the inconclusive empirical evidence for its effectiveness would be to hypothesize an inconsistency between stimulants' perceived and actual effects on healthy cognition. Specifically, people may use stimulants for cognitive enhancement because they feel that the drugs improve their performance, even in the absence of actual effects. For this to be the case, two conditions need to be satisfied: first, participants must perceive their own performance as higher; second, they must attribute this higher performance to the drug.

Addressing the former condition, a number of studies have asked whether self-estimation of performance increases as a function of stimulants. This idea was first considered by researchers in the middle of the 20th century, motivated in part by concerns about amphetamine's effect on the judgment of military personnel. For example, Davis (1947) summarized his experience with British soldiers in World War II by writing that "the subject who has taken amphetamine usually judges the effects more

favorably than the experimenter." Experimental evidence has provided converging support for this finding (Smith & Beecher, 1964; Hurst, Weidner & Radlow, 1967, despite a null finding in Baranski et al., 1997). In Smith & Beecher's double-blind, placebo-controlled trial on amphetamine, participants took a calculus test. Although they overestimated their performance in both conditions, the magnitude of overestimation was significantly greater in the amphetamine group. In a recent study with modafinil, a nontraditional stimulant, Baranski, Pigeau, Dinich & Jacobs (2004) reported a trend toward more positive evaluation of performance with modafinil compared to placebo in a battery of cognitive tests. The idea that drug effects on subjective assessment of performance may interfere with our ability to judge drug effectiveness for cognitive enhancement was raised more recently by Hall and Lucke (2010) who pointed out that, when taken by healthy people, stimulants may inflate self-confidence, while failing to improve actual performance. Although previous research has reported some evidence for amphetamine's effects on self-evaluation, no research study, to our knowledge, has assessed whether participants specifically attribute this improved performance to the drug. Only if this is the case can the subjective drug effects explain the growing stimulant enhancement use in the absence of firm evidence for actual effects. For this reason, in addition to measuring the objective effects of the MAS on cognitive performance, we also obtained rating of subjective perceptions of the effects of the ingested pills.

The Present Study

The purpose of the present double-blind, placebo-controlled, crossover study was to examine the actual and perceived cognitive enhancing effects of MAS on healthy young adults who were not sleep-deprived. MAS is equivalent to the brand name drug Adderall, which has been characterized as the "drug of choice" for cognitive

enhancement among college students (DeSantis, Noar & Webb, 2009). We predicted that, relative to placebo, Adderall would improve performance on a wide range of cognitive functions, including episodic and working memory, inhibitory control and creativity, as well as performance on tasks based on standardized tests. We further expected that low cognitive performers, as well as carriers of the *val-val* variant of the COMT gene would benefit from the drug more that high performers and *met-met* carriers, respectively. An alternative hypothesis was that MAS might evoke a subjective perception of enhancement, more salient than the drug's actual enhancing effects. If substantiated, either of these predictions would provide a possible explanation of the growing psychostimulant use among healthy people.

Method

Participants

Participants were 46 Caucasian native English speakers (22 male and 24 female), aged 21-30 (*M* age = 24, *SD* = 2.88), who responded to advertisements posted in the area of Drexel University and the University of Pennsylvania, as well as to email announcements at the University of Pennsylvania), inviting participation in tests of memory, creativity, intelligence and personality. Participants were excluded if they reported a history of medical conditions contraindicated for stimulant use, including any neurological or psychiatric disease, seizure disorder, high blood pressure, glaucoma, gastrointestinal blockage, heart disease, or thyroid problems. Also excluded were participants using any other stimulants or substances that could interact with amphetamine, including addictive, psychoactive, neurological and blood-pressure drugs; anti-histamines; non-prescription dietary supplements; weight-loss pills, and caffeine consumption estimated to exceed 700mg/day. Off-drug blood pressure measured to exceed 140/90 at the beginning of the study was another exclusion

criterion because of the likelihood that the drug would increase blood pressure further. Women who were pregnant or likely to become pregnant were not allowed to participate. We also excluded potential participants who had previously used psychostimulant drugs to rule out sensitization (Paulson & Robinson, 1995) as an explanation for enhancement effects and tolerance (Schenk & Partridge, 1997) as an explanation for a lack of such effects.

Drug

20 mg of mixed amphetamine salts (sulfate salts of dextroamphetamine and amphetamine, with the dextro isomer of amphetamine saccharate and d, l-amphetamine aspartate monohydrate, with d-amphetamine and l-amphetamine in 3:1 ratio) and placebo were administered in visually indistinguishable capsules. The selected dose is within the range of doses used in the enhancement literature, with some studies using lower doses, some higher, and some equivalent (Smith & Farah, 2011). The test drug was supplied by the University of Pennsylvania Investigational Drug Service.

Tasks

There were three versions of each task, to avoid repetition of items or trial orders across baseline, placebo and MAS conditions. The three versions were of moderate and comparable difficulty as determined by pilot testing. One version was consistently used for the baseline condition while the other two versions were used equally often in the MAS and placebo conditions. The 13 measures are listed in Table 1.1 and are described here.

Memory. The consolidation of information into memory is central to learning and hence to academic and life success. Two tasks assessed memory with measures of verbal and visual recognition and verbal recall.

Face Memory. In this test of episodic memory participants saw a sequence of 20 briefly flashed face images. Each stimulus was presented for 2250 ms, with an inter-trial interval of 750 ms. Encoding was followed by approximately two hours of cognitive testing, after which participants completed a recognition test. This test consisted of the 20 previously presented targets intermixed with 20 new faces. Presentation duration and inter-trial interval at test were the same as those at the encoding phase. Our main dependent measure was number correct (total number correct out of total number of trials presented).

Word Memory **(two measures).** Another test of episodic memory, this task presented participants with 25 words presented for 3s each with no intertrial interval. After approximately two hours performing other cognitive tasks, two measures of word memory were then obtained. In *word recall,* participants freely recalled as many words as possible. Performance was measured as number recalled. A *word recognition* test followed, in which participants viewed the 25 earlier words intermixed with 25 new words, presented in the same way as during encoding. The dependent measure in this case was number correct.

Working memory*.* Working memory is an aspect of executive function that involves active short-term maintenance of information and is essential for many forms of thinking and problem-solving. Two tasks assessed verbal and visual working memory.

Digit Span Forward and Backward **(two measures).** In this test of working memory, derived from a subtest of the Weschler Adult Intelligence Scale, the experimenter read digit sequences at a rate of approximately 1 digit per second and participants typed each sequence immediately after hearing it. In the Digit Span Forward task, 14 sequences were presented, with two each of lengths from 2 to 8.

Participants reported the digits in the order they heard them. In the Digit Span Backward condition, 14 different sequences of length 2-8 were presented and participants reported the digits in reverse order. The digit sequences gradually increased in length (from two to eight digits). A response was counted as correct only if all the digits within a sequence were reported correctly and in the correct order (i.e., no partial credit). Number correct for Digit Span Forward is generally viewed as a simple measure of maintenance capacity and Digit Span Backward as a measure of ability to simultaneously maintain and process information (The Psychological Corporation, 2002).

Object-2-Back. Object-2-Back tests the ability to maintain and update information in working memory despite interference (Postle, D'Esposito, & Corkin, 2003). Participants saw a sequence of 155 random polygons, each flashed briefly, for duration of 1000ms. Interstimulus interval was 500 ms. Participants had to press a button every time the currently presented object matched the object two shapes back. The dependent variable was the number of omission errors, which is associated specifically with working memory ability in n-back tasks (Oberauer, 2005).

Inhibitory control. The ability to withhold a habitual response or resist distraction by a salient stimulus is important for enabling us to act appropriately in many contexts. Two tasks assessed inhibitory control over responses and stimuli.

Go/No-go. Go/No-Go is a test of inhibitory control (see Braver, Barch, Gray, Molfese & Snyder, 2001), in which participants viewed a sequence of briefly flashed digits 1-9 (stimulus duration: 300ms; interstimulus interval: 400ms). Participants were asked to press a key as quickly as possible in response to all digits except for the digit 4. The digit 4 appeared on 15% of the 200 trials. For this task, the only opportunities for inhibitory control failure occur on the "4" trials and therefore the dependent measure

was the number of commission errors (i.e., the number of trials on which participants failed to withold a response to the “4”; Helmers et al., 1995).

***Flanker*.** This test of inhibitory control (Eriksen & Eriksen, 1974) presented participants with 200 images of five horizontally aligned arrows. Participants were instructed to indicate as quickly and accurately as possible the direction (left or right) of the central arrow. In congruent trials, all arrows pointed in the same direction. In incongruent trials, the middle arrow pointed in a direction opposite to that of the peripheral arrows. The sequence consisted of an equal number of congruent and incongruent stimuli. Each stimulus remained on the screen until the participant responded; the response initiated a 1s blank screen before the next stimulus appeared. The measure of inhibitory control, termed here inhibition cost, was the ratio of the median reaction time of incongruent trials to the median reaction time in response to congruent trials.

Creativity. Creativity is often defined as the ability to recombine familiar concepts in new and useful ways. It has been operationalized with tasks that require participants to find associations among disparate concepts and to view complex visual patterns in alternative ways. The two tasks used here were previously used by us in a study of the effects of MAS on creativity (Farah, Haimm, Sankoorikal, Smith & Chatterjee, 2008).

Remote Associations Test. In this test of convergent creativity (Mednick, 1962) participants must generate the word which associates a group of three other words. For example, presented with the stimulus triad “round – manners – tennis,” they had to answer “table.” The test included 15 triads, for each of which participants had 30 s to respond. The dependent measure was number correct.

Group Embedded Figures Task. Another measure of convergent creativity (Noppe & Gallagher, 1996) this test presented participants with complex geometric designs and a smaller element from the design. Within each larger design participants had to find and trace the specificed target element, which is "embedded," that is, not immediately apparent, given the overall visual gestalt of the design. An example is shown in Figure 1.1. Participants completed 6 items within a 2-minute time-limit for the whole test. The dependent measure was number correct.

Standardized tests. Among the many standardized tests of intelligence and achievement are Raven's Advanced Progressive Matrices, a nonverbal test of fluid intelligence, and the Scholastic Achievement Test (SAT), taken by college applicants in the US. The two tasks here were composed of individual items taken from these standardized tests.

Raven's Advanced Progressive Matrices. In this test of nonverbal intelligence (Raven, 1976) participants saw a series of abstract patterns, each of which had a missing piece. Participants had to choose the best fitting piece from 6 options. Each version of the test consisted of 12 items. Completion time for the whole test was limited to 10 min. The measure of interest was number correct.

Scholastic Achievement Test (SAT;* two measures*). This standardized test includes sections assessing "critical reading," "writing" and "mathematics." We selected questions from a book of practice tests and grouped them into two sections, "Verbal" and "Math." The former consisted of 48 multiple-choice questions completed under a 40-minute time limit. Question types (and corresponding number of questions) were as follows: Sentence Completion (7), Reading Comprehension (26), Improving Sentences (9), Identifying Sentence Errors (6). The Math section consisted of 27 questions (19 in multiple-choice and 8 in free-response format) testing algebra, geometry and other

miscellaneous high-school-level problems, to be completed under a 28-minute time limit without the use of a calculator. The measures of interest for *SAT-Verbal* and *SAT-Math* were number correct.

Perceived drug effect. Perceived effect was examined through the following self-report prompt: “The following question refers to all tests completed TODAY. How and how much did the drug influence (either positively or negatively) your performance on the tests? Please use the scale below. You answer can be any number between 1 and 100." The scale referred to was a line, ranging from 1 to 100, and labeled as follows: 1 = “the drug impaired my performance extremely”; 25 = “the drug somewhat impaired my performance”; 50 = “the drug had no effect”; 75 = “the drug somewhat improved my performance”; 100 = “the drug improved my performance extremely.

Procedure

The study took place over seven sessions, which included consent and practice (Session 1), followed by the full battery of cognitive tasks, for the baseline (i.e., no pill), placebo and MAS conditions (Sessions 2-7). Baseline testing (Sessions 2-3) always preceded drug/placebo testing (Sessions 4-7) to minimize the influence of practice effects on data from the placebo and MAS conditions. During on-pill Sessions 4-7, the order of drug administration was counterbalanced, in a way that 24 participants received MAS in Sessions 3 and 4, and 22 participants received the drug in Sessions 6 and 7 The timeline of the study, including session sequence and timing, is presented in Figure 1.2.

Session 1: Intake interview, instructions and practice. The first session consisted of consent procedure, followed by practice versions of the actual tests. The practice tests were identical to the experimental versions, except for comprising of fewer trials and different items. At the end of the session, participants were instructed to

abstain, for the rest of the study, from drugs containing stimulants or interacting with stimulants (or to notify the study personnel if they had to take such drugs). Participants were also asked to avoid heavy meals on test days.

Session 2 and 3: Baseline testing. Sessions 2 and 3 provided a measure of unmedicated (off-drug, off-placebo) performance. The placebo condition was not used as a measure of baseline, so that regression to the mean would not be mistaken for moderation by cognitive ability.

After the initial blood pressure measurement, participants completed an SAT test (one Verbal and one Math section) and a battery of cognitive tests (described above), respectively in Session 2 and 3. These baseline tests were a version of the to-be-administered on-pill battery.

Sessions 4-7: Testing on drug and placebo. The goal of these four sessions was to measure participants' cognitive performance on MAS and on placebo. At the onset of these sessions, participants reported the amount of sleep and caffeine consumption during the previous 24 hours. They answered questions on their diet and medication intake to determine compliance with earlier instructions and had their blood pressure measured at the beginning and end of these sessions. Participants also underwent a urine drug test to corroborate self report and deter use of excluded drugs (amphetamine, cocaine, barbiturates, benzodiazepine, phencyclidine and tetrahydrocannabinol). Female participants were administered a urine pregnancy test. Participants with positive results on any of these tests were excluded. After an initial blood pressure measurement (participants were excluded if the measurement exceeded 140/90), participants were randomly assigned to take either MAS (20mg) or a visually indistinguishable placebo capsule in a double-blind manner. A 75-minute waiting period followed. We chose this interval to ensure that the peak drug plasma

level, which is reached 2-3 hours after administration (Angrist et al., 1987) would occur during the testing. During the waiting time participants remained in the testing area and either read student periodicals or watched documentary DVDs (no homework or exciting movies were allowed). Five minutes before testing (70 minutes after drug intake), blood pressure was taken again and participants were excluded if the measurement exceeded 150/100. In sessions 4 and 6 the battery of tests included personality, mood and attributional style questionnaires (not relevant to cognitive enhancement and therefore not discussed further here) and test materials assessing verbal and mathematical abilities from the SAT. In sessions 5 and 7 the remaining cognitive tests were administered, in the same order for all participants. The two nonbaseline versions of the tasks were counterbalanced with both drug condition (MAS, placebo) and session order. After the cognitive battery (sessions 5 and 7) participants reported their perception of the pill's influence on their performance using the scale described earlier. At the end of the session, participants were reminded of the restriction on caffeine use and heavy-meal consumption for the rest of the study. If finishing the study, participants were thanked and paid.

Data Analysis Approach

Outlier removal. We removed outliers by excluding individual scores 3 *SD* above or below the mean of either the drug, placebo or baseline on each cognitive and subjective measure. If, on a particular measure, an individual participant's baseline, drug, or placebo score met the criterion for an outlier, we excluded all the data (i.e., MAS, Placebo *and* baseline) of this participant from analyses of that same task. This led to the exclusion of a total of 22 data points, or .85% of all data.

Missing Data. 143 task performance measures (or 5.55% of all data) were missing due to technical problems (50 data points), evidence of participants' failure to understand the task instructions (9 data points), or experimenter error (84 data points).

Statistical Tests. In overview, our approach to statistical testing was based primarily on mixed model analysis of variance (ANOVA) with drug (MAS or placebo) as a within-subjects factor and drug order and test version order as between-subject factors. Multivariate analysis of variance (MANOVA) was used to assess the effects of drug across all 13 measures of cognitive ability and ANOVA to assess drug effects on each individual measure of cognitive ability and on ratings of perceived enhancement. Moderation of cognitive enhancement by baseline ability and COMT genotype was tested within the same framework. We also use multivariate regression and simple bivariate correlation in order to test two specific relations involving non-categorical factors (the moderating effect of COMT *val* load and the relation between perceived and actual enhancement). All analyses were conducted in SPSS 20. The significance threshold was set to the standard cutoff of .05. Results are reported without correction for multiple comparisons, a lenient approach that maximizes our ability to identify positive results at the risk of increasing possible false positive results.

Results

Effects of Mixed Amphetamine Salts

Table 1.2 shows the means and standard deviations of performance in each task for the baseline, placebo and MAS conditions. To examine whether cognitive performance differed between MAS and Placebo sessions, we conducted a 2(Drug: MAS; Placebo) x 2 (Drug Order: MAS first; Placebo first) x 2 (Test Version Order: Version 1 first; Version 2 first) mixed-model MANOVA with repeated measures on the

first factor. The dependent variables were scores for 13 measures (listed in Table 1.1).On this test, the difference between MAS and placebo performance did not reach significance (F (13, 13) = 1.71, p = .17), indicating that there was no overall enhancing effect of MAS on cognitive performance in the tasks. We also failed to observe any significant two-way interaction between the drug conditions and either the drug order or the version order: (F (13, 13) = .59, p = .83; F (13, 13) = 1.28, p = .33, respectively). The absence of a Drug x Drug Order interaction indicates that MAS is no more or less enhancing when taken before or after the placebo session. However, given the inclusion of a Baseline condition before all MAS and placebo conditions, these results do not rule out the possibility that MAS could enhance performance with novel tasks. A marginally significant three-way (Drug x Drug Order x Version Order) interaction was observed (F (13, 13) = 2.14, p = .09). This interaction, which indicates differential drug effects on different versions of the tasks depending upon the order in which they were performed, does not lend itself to any obvious interpretation. The possibility that the versions differed in difficulty, and the order in which they were encountered synergistically compounded these difficulty differences, is not supported by a comparison of performance across versions from placebo conditions.

Although we began by testing the multivariate hypothesis that MAS would enhance overall performance across tasks, we also had a priori hypotheses about MAS effects on each of the 13 measures obtained in the project. We therefore followed up the MANOVA with a series of univariate 2 (Drug: MAS; Placebo) x 2 (Drug Order) x 2 (Test Version Order) mixed-model ANOVAs with repeated measures on the first factor. These analyses tested the effect of MAS on each of the 13 cognitive measures listed in Table 1.1. Again, these analyses revealed no effects of MAS and no two-way interactions between drug and order or drug and version for any of the 13 measures.

The three-way interaction trend noted above emerged as significant (without correction for multiple comparisons) for five of the thirteen measures: Face Recognition, Flanker, Remote Associations, Embedded Figures, SAT Math. All main effects and interactions are shown in Table 1.1.

Faced with null results for the effect of MAS on cognitive performance in these tasks, we asked whether differences in participants' sleep prior to the MAS and placebo test days could have obscured the drug's enhancing effect. Self-reported sleep duration did not differ significantly between the sessions (t (41) = .91, p = .37 for neurocognitive testing sessions; t (45) = .74, p = .47 for SAT sessions), and showed a trend in the opposite direction to that hypothesized here, toward more sleep before MAS test sessions (M = 7.12, SD = 1.26 for neurocognitive testing sessions; M = 7.15, SD = 1.30 for SAT sessions) than placebo (M = 6.89 h, SD = 1.53 for neurocognitive testing sessions; M =6.98, SD = 1.45 for SAT testing sessions).

Moderation of MAS Effect by Baseline Performance

To determine whether MAS enhances cognition for some people, with an effect that is moderated by baseline cognitive performance, we first separated participants into two groups according to whether their baseline performance was above or below the median and then conducted a series of 2 (Drug: MAS; Placebo) x 2 (Baseline Performance: Below-Median; Above-Median) x 2 (Drug Order) x 2 (Test Version Order) mixed-model ANOVAs with repeated measures on the first factor for each of the 13 cognitive performance measures. (MANOVA was not carried out because different participants fall in the upper and lower groups for different measures). Significant interactions between drug and baseline performance emerged on two measures: Word Recall (F (1,36) = 4.78, p = .04) and, replicating our earlier study of MAS effects on this task (Farah et al., 2008), Embedded Figures (F (1, 28) = 8.48, p < .01). There was also

a marginal trend toward significance for Raven's Progressive Matrices (F (1,29) = 2.83, p = .10). In all three cases, the pattern of means was consistent with the prediction, based on the literature discussed earlier, of relatively more enhancement for the lower performing participants. As shown in Figure 1.3, MAS tended to improve performance for the below-median baseline performers, while acting in the opposite direction for the above-median performers.

In addition to comparing the effects of MAS between higher and lower performing participants, we can also ask whether low performers, the subgroup exclusively expected to benefit from the drug, shows enhancement. This question was addressed by a series of 2 (Drug: MAS; Placebo) x 2 (Drug Order: MAS first; Placebo first) x 2 (Test Version Order: Version 1 first; Version 2 first) mixed-model ANOVAs with repeated measures on the first factor for each of the 13 measures only among the subsample performing below the median on baseline. The effect of the drug was significant on Word Recall (F (1, 16) = 6.71, p = .02, Embedded Figures (F (1, 12) = 8.41, p = .01; and Raven's Progressive Matrices (F (1, 16) = 5.36, p = .03). None of the remaining measures showed evidence of enhancement for the lower performing participants. See Table 1.3. for other results which, because extraneous to our prediction, are not discussed further.

Moderation of MAS Effect by COMT Genotype.

Given the findings reviewed earlier of moderation of amphetamine enhancement effects by COMT genotype, we divided participants into three groups depending on whether they had *val-val, val-met*, or *met-met* alleles of COMT. Because a MANOVA using genotype as a 3-level factor would not capture the ordering among the three groups we instead employed three alternative sets of analyses.

First, we conducted a regression analysis, which included *val* load, drug order and version order to predict a composite of the differences between MAS and placebo. The overall model was not significant, p = 0.28. Second, we carried out thirteen additional regressions to examine the effects of *val* load, drug order and version order on MAS effect (i.e., drug minus placebo score) on each separate measure. Overall regression models for SAT Math and Verbal were marginally significant (F (3,43) = 2.52, p = .07; F (3,43) = 2.25, p = .10, respectively). On SAT Math, the effect of COMT was significant (b = .35, t = 2.38, p = .02); this effect was near significant on SAT Verbal (b = .24, t = 1.64, p = .11). The patterns of means complied with the prediction that people with *val-val* genotype are more susceptible to enhancement than those with *met-met* genotype (see Figure 1.4).

Third, we used MANOVA to contrast the effects of MAS on the two groups of homozygous participants with a 2 (Drug: MAS; Placebo) x 2 (COMT genotype: *val-val; met-met)* x 2 (Drug Order) x 2 (Test Version Order) mixed-model MANOVA, as well as with corresponding ANOVAs for each of the 13 measures. Neither the multivariate test for COMT moderation was significant, (F (5, 1) = .64, p = .73), nor any of the univariate tests, p > 0.26 in all cases, with the exception of a significant Drug x COMT interaction on SAT Math (F (1, 11) = 13.06, p < .00) which again complied with the predicted pattern of relatively greater enhancement for homozygous *val* than homozygous *met* participants (see Figure 1.4). Additionally, a significant drug effect emerged on SAT Math: F (1, 11) = 5.63, p = .04, although as shown Figure 1.4, this main effect of drug was an overall impairing effect. Main effects of COMT genotype were found for Word Recognition (F (1, 10) = 5.42, p = .04), along with borderline significant effects for Word Recall (F (1, 10) = 4.65, p = .06), SAT Verbal (F (1, 11) = 4.24, p = .06) and Object-2-

Back Omissions (F (1, 10) = 4.22, p = .07). The *met-met* genotype was associated with better performance than the *val-val* in all three cases.

As with the analyses of baseline performance moderation, we followed up these analyses of genotype moderation with direct comparisons of drug and placebo performance in just the subjects for whom the drug would be expected, a priori, to be more helpful. We first carried out a 2(Drug: MAS; Placebo) x 2 (Drug Order: MAS first; Placebo first) x 2 (Test Version Order: Version 1 first; Version 2 first) mixed-model MANOVA with repeated measures on the first factor and the 13 performance measures as the dependent variables. The effect of MAS did not reach significance (F (2, 1) = .27, p = .81), nor did other effects or interactions, p > .49 in all cases. We then ran a series of 2 (Drug: MAS; Placebo) x 2 (Drug Order: MAS first; Placebo first) x 2 (Test Version Order: Version 1 first; Version 2 first) mixed-model ANOVAs with repeated measures on the first factor for each of the 13 measures only among the subsample homozygous for the *val* allele. The effect of the drug did not reach significance on any of the measures (all p's > .24). See Table 1.3. for other results which, because extraneous to our prediction, are not discussed further.

In sum, as with the analysis of moderation by baseline performance, we found mixed evidence for the moderation by COMT: little evidence supported the predicted moderation but when such moderation was observed, it was generally consistent with the hypothesis of relatively greater enhancement in carriers of the val allele.

Perceived enhancement

We examined MAS's effect on perceived enhancement through a 2 (Drug: MAS; Placebo) x2 (Drug Order) x2 (Test Version Order) ANOVA, with repeated measures on the first factor. A main effect of drug (F (1, 40) = 4.09, p = .05) indicated that

participants perceived MAS (*M* = 55.18, *SD* = 14.87) as slightly more beneficial for cognitive performance than placebo (*M* = 50.25, *SD* = 3.95).

Although the earlier analyses demonstrate that MAS did not enhance cognition by any of the measures examined, it is nevertheless possible that subjective perceptions of enhancement are related to degree of true enhancement. To test this, we correlated the difference in perceived enhancement between MAS and placebo, on the one hand, and the corresponding difference scores on each of our 11 cognitive performance measures (no measure of perceived enhancement was administered during SAT Math and Verbal sessions). Ten of the 11 correlations did not reach statistical significance. A significant correlation emerged between perceived and actual enhancement on Go/No-go (r = .33, p = .04). To assess the relation between subjective perceptions and performance on all of the tasks together, we also created a composite of the differences between MAS and placebo sessions from each measure. This composite score did not correlate significantly with perceived enhancement: r = -.06, p = .76. In sum, on average participants believed that the MAS had enhanced their cognitive performance more than placebo. This perception stands in contrast to the reality: There was no actual enhancement on average nor were participants who felt more enhanced by the MAS more likely to show a true enhancement effect.

Discussion

Conclusions and relation to wider enhancement literature

Does MAS enhance cognition in healthy young adults? Our study was designed to overcome several challenges that have hampered previous attempts to answer this question. It had sufficient power to detect a medium-size effect for any one measure of cognitive performance. We nevertheless failed to find enhancement with any of the 13

measures we used. Of course, a different drug or a different does of MAS have led to a different finding. Nevertheless, we can state that a standard clinical dose of a drug that is commonly used for cognitive enhancement did not enhance cognition in an adequately powered study. The most straightforward interpretation of these results is that MAS is not a powerful cognitive enhancer. If it does enhance cognition in healthy and adequately-rested young adults, the effects are likely to be small.

These findings raise the question of why many published studies find large effects of amphetamine on cognitive performance with tests of memory, executive function and other cognitive processes. We believe that the answer is related to a set of problems, specifically low study power, flexibility in specific outcomes to be tested and publication bias against null results, which bedevil all branches of science, as explained in Ioannidis's (2005) provocatively titled article, "Why most published research findings are false." The impact of these problems on psychology and neuroscience research, in the absence of any intentional malfeasance has been discussed by Ioannidis (2011), Lehrer (2010) and Simmons, Nelson & Simonsohn (2011) among others. Research on cognitive enhancement is not particularly susceptible to these problems, compared to other research topics, but neither is it immune to them. As a result, it is difficult to estimate the true robustness and effect size of cognitive enhancement with MAS and other stimulant medications by surveying the published literature.

On the assumption that the enhancing effects are real but are too small to be reliably captured in studies with sample sizes in the range typically used, one would expect a mix of positive and null results to be obtained. Of course, those effects that are found would show relatively large effect sizes, because only those results that by chance err on the large side would achieve significance. This is the pattern that we

have seen in the literature, particularly regarding the effects of amphetamine on executive functions (Smith & Farah, 2011).

In the present study we also tested the hypotheses that MAS is enhancing for subsets of healthy young adults, specifically those who are less cognitively capable or who are homozygous for the Val allele of the COMT gene. Here too we generally failed to support these hypotheses, although a minority of specific statistical tests showed the predicted patterns.

Finally, we found a small but reliable effect of MAS on judgments of enhancement, reminiscent of Davis's (1947) observations of soldiers in World War II quoted earlier. Participants believed themselves to be more enhanced by the pill when given MAS compared to placebo. Although not apparent for every individual participant, the overall tendency was for participants to feel that their cognitive performance has been enhanced by MAS. This may in part explain MAS's popularity as a cognitive enhancer.

Limitations of the present study

The present study was carefully designed to sample a wide array of cognitive abilities, to have adequate power and to measure potential moderators. In other respects, however, its design leaves some important questions unanswered. Most importantly, like most published studies in the enhancement literature (Smith & Farah, 2011), we did not vary drug dose and cannot know whether a higher or lower dose of the drug might have produced different results. Similarly, we did not test the cognitive enhancing potential of other enhancers, such as methylphenidate and modafinil, leaving open the possibility that these drugs may significantly improve healthy people's cognitive performance. We did not measure bioavailability of the drug (e.g., plasma amphetamine) and so cannot quantify how this varied across participants and sessions,

for example as a function of individual differences in drug metabolism or food consumed before a session. Different or more frequent assessments of the perceived effects of MAS might have revealed more nuanced results or measured perceived enhancement more reliably. Our participants were not representative of the general population; in addition to the restricted age range, they met a number of health and lifestyle criteria for inclusion, including never having used prescription stimulants and being low or moderate consumers of caffeine. Perhaps different results would have been obtained with people who have already self-selected to use stimulants or who enjoy large daily doses of caffeine.

Implications for neuroethics

The present results have several implications for the neuroethics of cognitive enhancement. We believe that the issues of fairness, freedom and agency, discussed so extensively in the neuroethics literature (e.g., Farah, Illes, Cook-Deegan, Gardner, Kandel, King, Parens, Sahakian, & Wolpe, 2004) are not moot despite the present results. It is of course true that the most thoughtful and incisive ethical analysis is pointless if applied to an inaccurate representation of the empirical facts of the matter. But we believe that Hall and Lucke (2010) are too dismissive of the realities of cognitive enhancement when they write "Guidelines for enhancement prescription are … premature. More skepticism needs to be expressed about neuroenhancement claims for pharmaceuticals and bioethicists should be much more cautious in … making proposals that will facilitate such use." (p. 2042). The present results suggest only that the effects of one currently available enhancement drug are small when measured in laboratory tests of memory, executive function and tests of intellectual aptitude. These results leave many questions unanswered.

Among the important open questions are: How helpful might a small enhancement effect be over time? Might the effects be larger when measured under real-world conditions (e.g., with distractions in the environment or for longer and hence more tedious tasks than the typical memory or executive function experiment) or in a different state (e.g., after sleep deprivation)? Does MAS exert a larger effect on other processes, such as motivation to work, which are not captured by laboratory studies of memory and executive function but which nevertheless impact academic and other cognitive work? Or are users primarily attracted to this drug because of the illusory perception of enhancement our participants reported? These are important questions for future research, which will furnish the needed empirical basis for discussions of enhancement ethics and policy.

CHAPTER 2

PRESCRIPTION STIMULANTS' EFFECTS ON HEALTHY INHIBITORY CONTROL, WORKING AND EPISODIC MEMORY: A META-ANALYSIS

The scientific and popular literatures both document the use of prescription medications by healthy young people to enhance cognitive performance in school and on the job (e.g., Smith & Farah, 2011; Talbot, 2009). This practice, called 'cognitive enhancement', has provoked wide discussion of its potential social, ethical and public health consequences (Farah et al., 2004; Greely et al., 2008; Sahakian & Morein-Zamir, 2011). Recently another question concerning cognitive enhancement has arisen: To what degree do the medications used for cognitive enhancement in fact improve the abilities of cognitively normal individuals?

In view of the prevalence of cognitive enhancement and the intensity of academic and policy interest in this practice, it is surprising that the answer to this question has not been clearly established. The empirical literature on the effects of these stimulants on cognition in normal subjects has yielded variable results, with some reviewers doubting their efficacy altogether. For example, in reviewing the literature on the cognitive effects of methylphenidate, Repantis, Schlattmann, Laisney & Heuser (2010) concluded that they were "not able to provide sufficient evidence of positive effects in healthy individuals from objective tests." Hall and Lucke (2010) stated "There is very weak evidence that putatively neuroenhancing pharmaceuticals in fact enhance cognitive function." Advokat (2010) concluded her review of the literature by stating that "studies in non-ADHD adults suggest that stimulants may actually impair performance on tasks that require adaptation, flexibility and planning."

Smith and Farah (2011) attempted to test the hypothesis that stimulants enhance cognitive performance in normal healthy subjects with a systematic literature review. We included studies of amphetamine and methylphenidate's effects on episodic and procedural memory and three categories of executive function: working memory, inhibitory control, and third category of other executive function tasks that did not fit into either of the first two. Results were mixed, large effects, small effects and null effects all reported. For example, over a third of the executive function studies reported null results. One interpretation of this pattern is that the drugs confer a small benefit, which may fail to be detected in some studies because of inadequate power. The other possibility that chance positive findings, combined with publication bias, may be responsible for the positive evidence that exists in the literature. Thus, despite the large literature included in our review, we were forced to conclude that "there remains great uncertainty regarding the size and robustness of these effects." Meta-analysis is a method that can distinguish between the competing interpretations of the findings in the cognitive enhancement literature.

The primary goal of the present meta-analysis is to obtain a quantitative estimate of the cognitive effects of the stimulants amphetamine and methylphenidate. They are commonly prescribed for the treatment of Attention Deficit Hyperactivity Disorder, but are frequently diverted for enhancement use by students and others (e.g., McCabe et al., 2006; Puolin et al., 2007; Wilens et al., 2008). Guided by the findings of Smith and Farah's (2011) review, we focus on the cognitive processes that seemed most likely to be enhanced by stimulants, specifically inhibitory control, working memory and episodic memory. In addition, because this earlier review found the strongest evidence of episodic memory enhancement after long delays between learning and test, we distinguish between episodic memory tested soon after learning (within 30 minutes following learning trials) and episodic memory tested after longer intervals (1 hour to 1 week).

The meta-analysis has two additional goals: One is to test hypotheses about moderators of the effects, that is, differences between studies that might account for the variability in effectiveness noted across different studies. For example, perhaps one of the stimulants is effective and the other less so, or perhaps low doses are more effective than higher doses. The final goal is to assess the role of publication bias in shaping the literature and potentially inflating effect size estimates. This would happen if, as hypothesized previously (e.g., Smith & Farah, 2011), underpowered studies obtained large statistically significant effects by chance and thereby entered the literature while the balancing effects of smaller or null results from similar studies remained unpublished.

Method

Search strategies

Online databases PubMed and PsychInfo were searched with key words "amphetamine" and "methylphenidate," each combined with each of the following: "executive function," "executive control," "cognitive control," "inhibitory control," "inhibition," "working memory," "flanker," "stop signal task," "stop task," "no-go," "card-sort," "ID/ED," "set shifting," "Sternberg memory," "Stroop," "Digit Span," "memory," "learning," "recall," "recognition," "retention." These searches were narrowed to exclude research on non-human subjects, qualitative studies, and non-empirical publications (e.g., review papers, meta-analyses, lectures, news articles, etc.). In addition, the reference sections of the following review papers were searched for relevant articles: Advokat (2010), Chamberlain et al. (2010), Repantis, et al. (2010) and Smith & Farah (2011). Finally, we searched the list of articles being reviewed by an American Academy of Neurology committee studying cognitive enhancement on which the last author serves. All research published through the end of December 2012 was eligible.

We also sought relevant unpublished data to include in the meta-analyses. Twenty researchers active in the area were contacted for unpublished data on amphetamine or methylphenidate effects on episodic memory, working memory or inhibitory control in healthy non-elderly adults. In addition, fourteen requests were made for additional data from studies published in the past 10 years but originally reporting insufficient data to calculate effect sizes. This led to obtaining 2 data sets of studies in progress or in submission, as well as additional effect size data from 4 published reports.

Criteria for study eligibility

Publication type and language. Empirical investigations in any report format were eligible for inclusion in the meta-analysis. These included journal articles, as well as dissertations, conference presentations and unpublished data sets. The latter three were considered in an attempt to minimize the influence of publication bias on the obtained effect size estimates. Only reports in English were included.

Participants. Eligible participants were young and middle-aged adults. Research on children, elderly, criminal or mentally ill participants was excluded. Studies were also excluded if the experimental procedure entailed sleep deprivation.

Research design: methodological quality. A double-blind, placebo-controlled design was required for inclusion. This criterion aimed to maximize the methodological quality of the meta-analyzed material.

Research design: intervention. Eligible interventions were orally administered amphetamine and methylphenidate, with drugs administered before the start of the cognitive protocol (e.g. not after learning in a memory experiment). We only included research on single dose administration (the only study on the effect of repeated

administration was excluded due to lack of consistency of intervention strength with the rest of the available research). In the included studies, the interval between drug administration and the cognitive task ranged between 30 min - 4.5 hours for amphetamine studies, and 40 min – 4.5 hours for research on methylphenidate. These intervals are within the medications' window of effectiveness (Vree & van Rossum, 1970; Angrist et al., 1987; Volkow et al., 1998). In addition, it is not unreasonable to suspect that these waiting times have ecological validity, with users working or studying similar intervals after drug intake.

Studies including multiple intervention arms, such as different drugs or transcranial magnetic stimulation (TMS), were included only if the effects of amphetamine and methylphenidate could be assessed in isolation (e.g. without concurrent TMS) and compared to placebo.

Cognitive systems under investigation. Four abilities central to academic and professional work were included, based on the findings of Smith and Farah's (2011) literature review. They were: *inhibitory control*, the ability to override dominant, habitual or automatic responses for the sake of implementing more adaptive, goal-directed behaviors; *working memory*, the capacity to temporarily store and manipulate information in the service of other ongoing cognitive functions; *episodic memory*, the ability to encode, store and retrieve task-relevant information, assessed shortly after learning (i.e. within 30 mins) and at longer delays (1 hour – 1 week). Whenever task descriptions were not sufficient to identify the cognitive function tested, the data were excluded.

Outcome measures. Performance can be measured by response time, overall accuracy, or specific types of error such as misses or false alarms. Overall in the literature, research reports varied in the types and number of outcome measures reported for each task. To maintain the validity and consistency of outcome measures in our analyses, we

designed an a-priori outcome selection procedure, shown in Table 2.1. Our outcome selection strategy favored the most widely used and construct valid measures, but also included second-best options, whenever our first choices were not reported. In general, we favored error measures over reaction time measures unless accuracy was near ceiling, in which case reaction time data, if available, were coded. On tests of inhibitory control, instead of overall accuracy, more specific accuracy measures (or the relationships thereof) were used, such as a measure of false alarms on Go/No-go tasks or the contrast in performance on incongruent and congruent trials of Flanker and Stroop. Whenever relevant, our main outcome measure was tailored to the specific design of the task. Particularly, two variants of the Stop Signal Task of inhibitory control have been used in the examined literature: a version where the probability of stopping is allowed to vary and is the main measure of inhibition (e.g., Fillmore et al., 2005), and a version where the probability of stopping is held constant (e.g., de Wit et al., 2002, Logan et al. 1997), in which case stop signal reaction time is the main outcome. Eligible outcome measures for each task are shown in Table 2.1.

Process of determining study eligibility

The search process, summarized in Figure 2.1, led to the identification of a total of 1799 titles, which were narrowed down to 1505 after 294 duplicate papers were removed. After screening the titles of these papers, additional 1304 reports were excluded for not meeting the inclusion criteria. The remaining 201 studies were assessed for eligibility by applying the exclusion criteria to the abstract, and, in case of insufficient data, to the full text.

Of the remaining 201 studies, 73 were excluded because the measured cognitive constructs (e.g., simple reaction time, sustained attention, creativity, intelligence, fear

conditioning, motor performance, reward processing, probabilistic learning, etc.) were outside the scope of the present review. Twelve studies failed to meet the criteria for eligible participants (mice: n = 1; elderly participants: n = 6; children: n = 2; mentally ill participants: n = 2, including 1 study on ADHD and 1 study on cocaine abuse; criminal participants: n = 1). Eighteen reports lacked a double-blind placebo-controlled design (when these design features were not explicitly mentioned, the study was excluded). 16 reports were excluded due to ineligible intervention. These included 4 studies which tested drugs other than amphetamine or methylphenidate; 4 studies in which drugs were administered intravenously; 4 studies conducted in the context of sleep deprivation; 2 studies in which outcomes were measured under TMS; 1 study in which drug administration followed (as opposed to preceding) learning; 1 study which tested the effect of multiple drug doses. Seven studies in language other than English were excluded. Four studies could not be retrieved from available online and paper sources. Four studies were excluded because duplicating the data of already included research. In 19 of the remaining otherwise eligible studies, reported data were insufficient to calculate effect size and authors did not respond to our requests for the needed additional information. The final analyses were based on 48 papers reporting at least one relevant effect size (44 published reports, 3 unpublished data sets and 1 dissertation with a total of 1409 participants). The first and the second author independently conducted the eligibility determination procedures; disagreements were resolved by consensus after reviewing the experimental reports.

Coding procedures

All studies were coded by the first author, according to a standardized coding manual. Coded variables included: means and standard deviations for performance under

drug and placebo; sample size; outcome measure; effect direction; significance level; and six moderators. The moderators, and rationale for examining their effects, were as follows.

1) Drug (methylphenidate vs. amphetamine). This moderator analysis was conducted to examine if amphetamine and methylphenidate differ in their cognitive enhancement potential. To our knowledge of the enhancement literature, no previous study has compared the enhancement effects of these two medications.

2) Dose (low vs. high). The cognitive effects of stimulants are dose-dependent (e.g., Robbins, 2000). In examining the role of dose in enhancement effects, we defined a "high" dose as amphetamine > = 20 mg and methylphenidate > = 40 mg. Doses below these benchmarks were coded as "low."

3) Caffeine restriction (present vs. absent). We explored the possibility that stimulants may be especially helpful in countering caffeine withdrawal, while possibly having limited effects on non-caffeine withdrawn individuals. The presence or absence of instructions to abstain from caffeinated beverages on the day of the experiment was coded as a possible moderator.

4) Gender distribution in the sample (percent male participants). In the past, higher rates of enhancement use have been reported among male students (e.g., Teter et al., 2005) and differences in stimulants' subjective effects have been shown to vary as a function of gender and menstrual phase (White, Justice & DeWit, 2002). The percentage of males in the study sample was therefore tested as a moderator.

5) Risk of ceiling or floor effects (suspected vs. not). Ceiling and floor effects could attenuate the estimated effect size. In these analyses, we examined whether the effect size in studies with no restriction of range differed from the effect size estimate in studies with

suspected floor or ceiling effects. A study was coded as being at risk of range restriction if the larger among the means in the drug and placebo conditions was less than 1 SD away from the scale's floor, or if the smaller mean was less than 1 SD away from the scale's ceiling. In case of moderation, our goal was to focus on the effect size estimate in the group of studies without suspected floor or ceiling effects.

6) Reason to publish if drug effects are null (present vs. absent). For the purpose of assessing publication bias for reports of behavioral effects of stimulants, we distinguished between effect sizes from studies that focused only on the effects of amphetamine or methylphenidate on healthy individuals and studies that also included clinical groups, other drugs or nonbehavioral measures such as PET, fMRI, EEG or ERP. We expected that smaller stimulant enhancement effects would be published in the context of studies addressing multiple questions (due to the higher likelihood of a positive finding given multiple measures and the greater resources invested in measuring neural activity and administering multiple interventions).

Effect sizes were calculated using means and standard deviations. Where these descriptives were not presented, we estimated them from published graphs. We favored descriptive over inferential statistics based on previous research showing that, in repeated-measures designs (the majority of the included studies), effect size estimates from descriptive statistics are less biased than those from repeated-measures inferential statistics (Dunlap, Cortina, Vaslow & Burke, 1996). In the absence of descriptive data, we estimated effect sizes from *F* (provided df = 1), *t* and/or *p*-values. If effect sizes were directly reported, we estimated their confidence intervals for $r_{equivalent}$ (Rosenthal & Rubin, 2003) and converted the values to *d*. When data were unavailable from either reports or from graphs, they were requested from authors.

The second author independently coded 44% of the means and standard deviations (including data, estimated from graphs) in the placebo and drug conditions. Analyses of reliability showed excellent agreement (two-way mixed-model ICC for absolute agreement > .99 in all cases).

Handling of missing data. Effect size data could not be retrieved or calculated from 19 reports. We performed all meta-analyses excluding all missing data. We did not impute data in missing cells because we had no reason to infer either zero or average sizes of these unreported effects (Cooper, 2010). In other words, we had no sufficient data to ensure that these analyses would improve our effect size estimates, instead of introducing error.

Statistical Methods

Effect size metrics. Hedge's g was used as the primary effect size measure, whereby a value of .2 is conventionally considered small, .5 is considered medium and .8 is considered large. Hedge's g is obtained by multiplying the effect size Cohen's *d* by a coefficient J which corrects for the tendency for studies with small sample sizes to bias the mean effect size positively due to publication bias: J = $1-\left(\frac{3}{4\times df-1}\right)$. In combining effect sizes, each was weighted by an estimate of its precision, i.e., the inverse of the squared standard error of the effect size.

For within-subjects designs, employed in the vast majority of the meta-analyzed papers, we have the option of calculating the effect size in two ways. Typically for such designs, a measure of performance *change* is scaled by units of variability of *change*. This addresses the question, "How much drug-related benefit can one expect, relative to the

variability of change scores in the sample?" Alternatively, the effect size can be expressed, as in with between-subjects designs, as the size of the drug treatment effect on performance, *measured in units of performance variability*. Specifically, using this approach, the difference in performance attributable to the drug is measured against the standard deviation of the sample's placebo performance. In effect, this addresses the question "how far along the distribution of normal performance does the drug push subjects?" This question is very appropriate to the study of cognitive enhancement when used to gain a competitive edge relative to an unmedicated population. Additionally, some authors have argued that "subject differences are always of theoretical interest" because "they are present in the population to which we want to generalize," justifying the calculation of effect sizes from either within- or between-subject designs in units of variability (Cortina & Nouri, 2000, p. 49). We report both types of effect size analysis here, placing primary emphasis on effect sizes measured relative to normal variability.

To conduct our primary analyses, we included research with both within- and between-subjects designs, and relied on effect sizes calculated as shown below. These formulas, typically used for between-subjects designs, were modified so that the observed standard deviations in the placebo condition are entered in the analyses as values for both medication and placebo conditions. In particular:

$$g = J \times \frac{M_{DRUG} - M_{PBO}}{SD_{POOLED}} \quad ,$$

where: $\mathrm{SD_{POOLED}} = \sqrt{\frac{((\mathrm{N_{DRUG}}\text{-}1)\mathrm{SD^2_{PBO}} + (\mathrm{N_{PBO}}\text{-}1)\mathrm{SD^2_{PBO}}}{\mathrm{N_{DRUG}} + \mathrm{N_{PBO}}\text{-}2}}$ and $\mathrm{SE} = \mathrm{SD_{POOLED}} \times \sqrt{\frac{1}{\mathrm{N_{DRUG}}} + \frac{1}{\mathrm{N_{PBO}}}}$

In these analyses, t, F and p values were used to derive effect sizes from between-subjects designs, using the following formulas: Hedges' $g = J \times \frac{t\sqrt{2}}{\sqrt{\frac{2N_1N_2}{N_1+N_2}}}$ and $SE = \sqrt{\frac{1}{N_1} + \frac{1}{N_2} + \frac{d^2}{2(N_1+N_2)}}$. Inferential statistics from within-subjects designs were not included in these analyses because they inherently reflect drug effects relative to variability of change, rather than relative to performance variability.

Our secondary analyses focused on the change score (drug minus placebo), specifically, the average benefit due to drug, relative to variability of change within the sample. Only within-subjects designs contained information relevant to this question. To calculate Hedge's g for change in within-subjects designs, the following formulas were used.

$$g = J \times \left(\frac{(M_{DRUG} - M_{PBO}) \times \sqrt{2(1-Corr)}}{SD_{DIFF}} \right),$$

where: $SD_{DIFF} = \sqrt{SD^2_{DRUG} + SD^2_{PBO} - 2Corr SD_{DRUG} SD_{PBO}}$, and $SE = \frac{SD_{DIFF}}{\sqrt{N}}$.

Alternatively, Hedges' g = $J \frac{t}{\sqrt{N}}$, and $SE = \frac{1}{\sqrt{N}} \times \sqrt{1 + \frac{d^2}{2}}$. These formulas require the value of the correlation between repeated measures, which were not reported in the published studies. These values, necessary to adjust for the dependency between repeated measures in effect size calculations, were estimated based on similar data sets[1],[2].

[1] Correlations were obtained from Ilieva et al., 2013 (Flanker, Go/No-go, NBack, Digit Span Backward and Forward, delayed memory for words and faces; 46 partipants), Mintzer et al., 2007 (NBack, Sternberg memory task, delayed memory for

Handling of studies with more than one effect size. One of the assumptions of meta-analysis is that each effect size comes from an independent sample. If this assumption is violated by the inclusion of more than one effect size per study, between-study variance will be underestimated and the significance of the summary effect size will be overestimated. The following steps were taken to reduce the available data to a single effect size per study.

1) When effect sizes for more than one construct per study were available, data on each construct (i.e., inhibition, working memory, short-term and long-term episodic memory) were separated in an individual meta-analysis.

2) When multiple doses of a drug were compared to placebo within the same study, effect size data from all doses were coded and averaged.

3) When, in a given study, effect sizes were reported for more than one eligible task and/or measure per construct, a single average effect size estimate per construct was obtained.

4) When outcome data were available from various time intervals after the administration of the drug (e.g., when inhibitory control was tested 1 hour, 2 hours and 3 hours after drug administration, or when long-term episodic memory was measured at various different retention intervals), the average effect size was entered in the main analyses.

---

words, 18 participants) and Hamidovic et al. (2009), combined with a set of unpublished data from Dr. Harriet de Wit's laboratory (Stop Signal task, 299 participants). When correlations for a given task (e.g., NBack) were available from more than one data set, we estimated a composite through meta-analyzing the available correlations based on a random effects model. For tasks for which data on observed correlations were lacking, we imputed an estimate of the correlation for the corresponding cognitive construct, obtained through meta-analyzing the available observed correlations for tasks within that construct based a random effects model.

2 To estimate the potential for error in case of inaccurate imputed correlations, we repeated our main analyses after varying the correlations between .2, .5 and .8. This led to minimal changes in the reported patterns of findings (largest change in effect size was g = 0.02).

Fixed vs. random effects model. A fixed effects model assumes that the only source of effect size variability is sampling error. It therefore produces an effect size estimate that describes the analyzed studies but cannot be generalized to other trials. By contrast, in a random effects model variability is assumed to arise from both sampling error and between-study variability. Effect sizes derived from this model can be generalized to research outside of the analyzed studies. For the present meta-analysis, we selected a random-effects model because of the variability between individual studies in each meta-analysis (different drugs, doses, waiting times between drug administration and testing, measures of each specific cognitive function, individual differences between samples), and also because we wanted to generalize the findings beyond the examined research.

Estimation of heterogeneity. Tests for heterogeneity determine whether the dispersion of the individual effect sizes around their mean value is greater than predicted solely on the basis of subject-level sampling error. One of the tests employed uses the Q statistic, which, if significant, rejects a null hypothesis of homogeneity. The second test, based on the I^2 statistic, generates an estimate of the between-study variance as a percentage of the total variance (between subjects plus subject-level). Conventions for low, moderate and high heterogeneity correspond to I^2 values of 25, 50 and 75 (Lipsey & Wilson, 2001).

Moderator analyses. Most commonly in the literature, moderator analyses are conducted only following a finding of significant heterogeneity. In contrast to this approach, we decided to conduct moderator analyses regardless of the results of the heterogeneity tests because a homogeneous set of findings may emerge either in the absence of moderators, or in the presence of moderators whose effects cancel each other out.

We examined the effect of the dichotomous moderators described earlier, using mixed effects analyses. This analytical model assumes that the effect size variation is due to a combination of systematic associations between moderators and effect sizes, random differences between studies and subject level sampling error. Finally, the moderating role of gender composition (measured as % male), was examined through meta-regression, given the continuous nature of this moderator.

A feature of the data on some moderator variables demanded the following modification in some of the analyses. When analyzing the moderating role of dose and ceiling/floor effects, there were a few cases of more than one level of the moderating variable for per study (this occurred more than one drug dose was administered per sample, or when floor/ceiling effects were suspected for one outcome within a study, but not for another). In these cases, we relied on two approaches to analysis. First, to satisfy the assumption of independence between effect sizes, we excluded studies which included data on more than one level of each moderator variable. In a second version of the analyses, we used the shifting-unit method of analysis (Cooper et al., 2010). The shifting-unit method allows violation of the assumption of meta-analysis in which a study can contribute an effect size to each level of the moderator (e.g., high and low dose). The advantage of the first approach is that the analysis assumptions remain unviolated; the advantage of the second approach is that it makes use of maximum possible data points. The findings based on the two approaches were in agreement, so we only report data based on the second one.

Publication bias. Publication bias refers to the greater tendency of studies with significant results to be published than non-significant findings. Publication bias can therefore bias the results of meta-analyses because the more significant findings typically

have larger effect sizes than those remaining in file drawers (Lipsey & Wilson, 2001). To minimize bias in the current meta-analysis, we made efforts to locate and retrieve unpublished data (see *Search Strategies* above). Additionally, we used three methods to assess the evidence for publication bias and the stability of the effect size estimates and to determine unbiased effect sizes: funnel plots, fail-safe N and trim and fill (Lipsey & Wilson, 2001). These analyses were conducted without correcting effect sizes by the factor J, described earlier. Only data from published reports were included in these analyses.

A *funnel plot* permits a qualitative test of publication bias, by showing the effect sizes of the analyzed studies plotted against an estimate of those studies' precision (the inverse of standard error of the effect size in our graphs). Effect size estimates from more accurate studies (towards the top of the graph) should cluster closely around the true effect size, while effect sizes from less accurate studies should appear more broadly dispersed below. In the absence of publication bias, the more broadly dispersed effect size estimates should extend in a roughly symmetrical arrangement to either side of the more accurate estimates. A negative skew, where points in the lower left quadrant appear to be missing, is consistent with the operation of publication bias.

In cases of publication bias, the *trim and fill* procedure calculates an unbiased estimate of the effect size. In this procedure, the most extreme positive effects are removed ("trimmed") from analysis and a mirror image of the trimmed effect sizes with the opposite direction is then imputed. Unbiased estimates of the overall effect size and its variance are calculated, respectively, from the trimmed and filled data.

The *fail-safe N* indicates the number of studies with a zero effect size that, if added to the analysis, would render the obtained mean effect size non-significant. The value of fail-safe N is considered large (and publication bias an unlikely influence on the effect size

estimate) if it exceeds 5k+ 10, where k is the number of meta-analyzed studies (Rothstein et al., 2006).

Tests for outliers. The presence of outlier effect sizes was assessed through the Sample-Adjusted Meta-Analytic Deviancy (SAMD) statistic. For each study, the value of this statistic represents the difference between this study's effect size and the point estimate of the effect size uninfluenced by this study, a difference weighed by the relevant variance terms (Huffcutt & Arthur, 1995). An effect size was considered an outlier if it met both of the following two criteria (Sockol, Epperson & Barber, 2011): First, in a scree plot of the distribution of absolute SAMD values, it deviates markedly from the slope (Huffcutt & Arthur, 1995). Second, it falls in the top or bottom 2.5% of the t distribution (which the SAMD distribution approximates). This conservative, two-pronged method for outlier detection was chosen because outliers could result from either error or true between-study variation (Sockol et al., 2011).

Software. The data were analyzed primarily using Comprehensive Meta-Analysis 2.0, with the exception of meta-regression analyses, completed in R 3.0.0.

Results

Overview of results. We report meta-analyses for the effects of stimulants on the four constructs of interest: inhibitory control, working memory, short-term episodic memory and delayed episodic memory. Two sets of results are presented, corresponding to the two different ways of measuring effect sizes from within-subjects designs described earlier. For each cognitive construct we first present meta-analyses of within- and between-subjects studies combined, measuring the size of the drug effect relative to variability in the normal

population. We then present the effect sizes estimated in separate meta-analyses for within-subjects and matched-groups studies using the formula for within-subject effect sizes described earlier. For the main analyses we also report the results of moderator analyses and three measures related to publication bias. In reporting our secondary analyses we do not detail the results of moderator and publication bias analyses, which in all cases were qualitatively similar to the results in our main analyses. Most effect sizes were small. Evidence of publication bias emerged in two cognitive domains. Characteristics of all effect sizes (outcomes, magnitude of effect, sample sizes, values of moderator variables) are presented in Tables 2.2-2.5.

Stimulants' effects on healthy inhibitory control

25 studies (including 2 unpublished) reported sufficient data to calculate the size of stimulants' effect on inhibitory control. After examining the values of the SAMD statistic, no value fell within the top 2.5% of the distribution or notably deviated from the relatively flat line of the scree plot. Not all of these studies were suitable for each analysis: i.e., a study whose effect size was derived from a repeated-measures t-value was excluded from analyses relative to normal variability; and a between-subjects study was excluded from analyses relative to variability of change. Data for calculating effect sizes relative to normal variability were available from 24 studies (see Table 2.2); effect size relative to variability of gain scores was also estimated from 24 studies.

Stimulants' mean effect on inhibitory control, when measured relative to normative variability of performance, was small but significantly different from zero: Hedges' g = 0.20, 95% CI [.11; .30]. Effect size measured relative to the variability of gain scores was similarly small and significantly different from zero: Hedges' g = 0.19, 95% CI [.11; .26]. No evidence for between-study heterogeneity emerged: Q (23) = 7.82, p > 0.99; I^2 = 0.00. Moderator

analyses indicated that none of the candidate moderators impacted significantly the stimulant effects on cognition (all p's > 0.20).

A funnel plot based only on the published studies (N = 22) showed no evidence for publication bias: the distribution of studies was roughly symmetrical (Fig. 2.2). The trim and fill procedure led to the exclusion of no study, and the adjusted effect size estimates remained the same as reported above. However, the fail-safe N method indicated that 39 studies (less than two studies per each published report) with an effect size of zero would nullify the obtained results. Taken together, the lack of negative skew in the funnel plot and the robustness of the effect-size estimate to trim-and-fill adjustment, converge to suggest that the effect estimate obtained for inhibitory control is most likely not affected by publication bias. In other words, there is no evidence to suspect that the relatively modest number of studies needed to nullify the result have remained in file drawers.

Stimulants' effects on healthy working memory

Effect size data on stimulants' effects on working memory were available from 23 studies, 3 of which were unpublished. None of the effect sizes were outliers by our criteria. Relevant statistics for calculating ES relative to normal variability were available from 20 studies (Table 2.3). Effect size relative to variability of gain scores was calculated based on 23 studies with within-subjects or matched-groups designs.

Our main analyses indicated a near-significant small stimulant effect on working memory: Hedges' g = 0.13, 95% CI [-0.02; 0.27]. When measured relative to variability of the gain scores, the effect size was again estimated to be small, but this time reached significance: g = 0.13, 95% CI [0.06; 0.20]. There was no significant evidence for

heterogeneity: Q (19) = 7.74, p = 0.99, I^2 = 0.00. Moderator analyses were performed, but no evidence emerged for moderation by any of the examined variables (all p's > 0.57).

The funnel plots, based on published studies only (Fig. 2.3) showed slight negative skew. The trim and fill procedure trimmed 5 data points, reducing the above-reported effect size to a non-significant trend of *d* = 0.06, 95% CI [-.08; .20]. Because the gain score effect size was significant, whereas the primary effect size was not, here we also report the trim-and-fill results from our secondary analyses, where the effect size was again reduced to non-significant: *d* = 0.06, 95% CI [-0.03, 0.15], given a negatively skewed funnel plot. Taken together, the trim-and-fill correction and the skew of the funnel plot, suggest the presence of publication bias. Fail-safe N analyses were obviated by the lack of significance in the obtained effect size estimate.

Stimulants' effects on healthy people's short-term episodic memory

14 effect sizes (1 unpublished) were considered for inclusion in the meta-analysis. Two SAMD values, equaling -8.53 (Burns, 1967) and 2.18 (Zeeuws et al., 2010a), exceeded the cutoff for exclusion and deviated markedly from the relatively flat line on the scree plot of absolute SAMD values. Therefore, these studies were excluded from further analyses after confirming correct data entry.

Based on 12 studies (see Table 2.4), the mean effect of stimulants on short-term episodic memory, relative to normal variation of performance, was small but significant: Hedges' g =0.20, 95% CI [.01; .38]. This was similar to the result observed when the effect size was measured relative to variability of gain scores (12 studies): Hedges' g =0 .22, 95% CI [.09; .35]. No evidence for heterogeneity emerged in our main analyses: Q (11) = 4.44, p

= 0.96, I^2 =0 .00. Moderator analyses indicated no significant influence of any of the examined moderators (all p's > 0.64).

A funnel plot, based on the 11 published studies, showed slight negative asymmetry (Fig. 2.4), despite the largest study having the largest effect. The trim and fill procedure trimmed 3 studies, reducing the effect size estimate to a non-significant d = 0.12, 95% CI [-0.06, 0.29]. According to the fail-safe N procedure, a mere 2 studies with an effect size of zero would be needed to nullify the obtained effect, casting doubt on the robustness of the effect.

Stimulant effects on healthy people's delayed episodic memory

12 effect sizes describing stimulants' effects on delayed episodic memory were reported. One outlier was excluded, given a SAMD value of 3.35 (Zeeuws et al., 2010a), which fell in the top 2.5% of the distribution of SAMD scores.

Based on the remaining 11 effect sizes, estimated relative to normal variability (see Table 2.5), stimulants' mean effect on delayed episodic memory was significantly different from zero and medium in size: g = .45, 95% CI = [.27, .63]. Similarly, analyses focusing on the mean gain, relative to the sample's variability of change, showed a medium-sized effect: Hedges' g =0.44, 95% CI [.26; .62]. There was no evidence for significant between-study heterogeneity: I^2 = 0.00, Q (10) = 9.67, p = 0.47. We found a small but significant moderating effect of gender: Q(1) = 7.44, p < 0.01, beta = 0.01, with larger drug effects for larger proportions of men in samples. In addition, there was a significant moderating effect of dose: Q (1) = 5.49, p = 0.02, indicating a larger effect for the smaller dose: g = 0.64, 95% CI [0.40; 0.88] than the larger dose: g = 0.20, 95% CI [-.08; .48]. Note that these moderation effects are confounded with each other and with research group: all studies

that used low doses of stimulants came from the same research group, tested only male subjects, and tended to test memory over longer retention intervals (1 hour – 1 week), while among tests of the high drug dose, the percent of men in the sample ranged between 48-70% and retention intervals, with one exception, were 2 hours. No other factors were found to significantly moderate stimulants' effects (all p's > 0.52).

The funnel plot of these studies was negatively skewed, suggesting publication bias (Fig. 2.5). The trim and fill method trimmed 5 studies, reducing the estimated effect size to d = 0.26, 95% CI [0.04; 0.47]. According to the fail-safe N procedure, 59 studies were needed to nullify the significance level of the result. The negative skew of the funnel plot, combined with the trim and fill correction, suggest the presence of publication bias and indicate that the true effect size may be small. It is important to note, though, that inferences from the funnel plot must be qualified by the presence of significant moderation (see Lau et al., 2006). In particular, the studies with the six largest effect sizes came from the same lab and tested the effect of a low stimulant dose on male-only samples, in part, over relatively longer retention intervals. Four of the five remaining studies with smaller effect sizes came from other research groups, and examined the effects of a high stimulant dose on a mixed-gender sample over relatively shorter delays. Thus, the funnel plot might reflect true publication bias, or might be driven by between-study differences. If the latter is the case, the trim-and-fill-adjusted effect size may be be underestimating the true effect size (e.g., Peters et al., 2007). Unfortunately, the proposed methods of unconfounding publication bias and moderating factors (e.g., conducting funnel plot analyses within a subgroup of studies) are applicable only to large meta-analyses (see Peters et al., 2010).

Discussion

Summary and interpretation of results

Earlier research has failed to distinguish whether stimulants' effects are small or non-existent (Smith & Farah, 2011; Ilieva, Boland & Farah, 2013). The present findings supported generally small effects of amphetamine and methylphenidate on executive function and memory. Specifically, in a set of experiments limited to high-quality designs, we found a small but significant degree of enhancement of inhibitory control and short-term episodic memory. Effects on working memory were small and significant in one of our two analyses. Delayed episodic memory was unique in showing a medium-sized effect. However, both working memory and delayed episodic memory findings were qualified by possible publication bias.

Several potentially important moderators were tested because of their scientific relevance for understanding the effects of stimulants on cognition and their practical relevance in determining whether stimulants might be more effective cognitive enhancers under some circumstances than others. Moderator analyses yielded only a few significant findings. Stimulant effects on delayed episodic memory were moderated by gender, with larger effects for samples with more males, and by dosage, with larger effects for smaller doses. Unfortunately, these two moderators were confounded in the studies analyzed, and also confounded with research laboratory and retention interval, so we cannot draw firm conclusions about the effects of gender or dose.

Where no effects of moderators were found, this may be due to uncertainty or imprecision in moderator coding, for instance the dichotomization of drug dose, or the

possibility of non-linear relationships between drug effect and the continuous moderators of sample gender and dose. Finally, partly for the sake of limiting the number of comparisons and partly due to limited availability of the relevant information, we examined only a subset of all relevant moderators. For instance, we did not explore the moderating role of participant age, level of education, waiting time between drug administration and testing, length of testing session or time of day. Moderators of great interest, which might be expected to affect results based on previous studies but which could not be assessed due to insufficient available data, include individuals' baseline cognitive ability and individuals' variants of dopamine-related genes such as COMT and DRD2 (see Mattay et al., 2003 and Hamidovic et al., 2009, but see also Wardle et al., 2013 and Ilieva et al., 2013 for null results of COMT's moderating effects). Consistent with the nonmonotonic relation between dopamine levels and performance, there is evidence that stimulants can impair performance in normal individuals who are especially high-performing (Farah, et al., 2008; De Wit, Crean & Richards, 2000; de Wit, Enggasser & Richards, 2002, Mattay et al., 2000). It remains possible that some individuals who would not qualify for a diagnosis of ADHD could nevertheless benefit from stimulants to a greater degree than indicated by the present results, and that some individuals could be impaired.

Neuroethical Implications

What do the results reported here imply for neuroethical issues surrounding the use of stimulants for enhancement? Should we be concerned about the fairness of students and workers competing with the help of stimulant drugs? Is there a genuine benefit to be weighed against the risks of using these prescription drugs for enhancement? The overall small effects of stimulants on healthy people's inhibitory control, working and episodic memory might be taken to mean that these drugs would not deliver a practically significant

performance advantage, and neuroethical discussions are therefore moot at best (and encouraging a false belief in the drugs' efficacy at worst, e.g., Hall & Lucke, 2010).

Nevertheless, the present findings provide reason to temper these and other more extreme skeptical assessments of stimulant medications for cognitive enhancement of healthy, cognitively normal individuals. Furthermore, small effects can make a difference in academic and professional outcomes. Even on a single occasion, a small effect might make the difference between good and very good performance, or between passing a school entrance or licensing exam or failing. It is also possible that these drugs may give a larger boost to cognitive functions not examined here (e.g., sustained attention, processing speed), to people not specifically studied in this meta-analysis (e.g., healthy participants with low cognitive performance or specific combination of genotypes), or to performance under conditions not tested here, for example fatigue, sleep deprivation, extreme distraction or repeated stimulant intake (e.g., Breitenstein et al., 2006). A final possibility is that prescription stimulants enhance work performance by altering users' emotions about, and interest in, tasks they would otherwise find boring and unrewarding (Vrecko, 2013; Ilieva & Farah, 2013).

The results of this meta-analysis cannot address these possibilities. Thus, there may well be solid reason for continued discourse on the effects, misuses, and ethical implications of cognitive enhancement with stimulants, a discourse, which the present data can importantly inform.

CHAPTER 3

ATTENTION, MOTIVATION AND STUDY HABITS IN USERS OF UNPRESCRIBED ADHD MEDICATION

Recent research has cast doubt on the cognitive enhancement potential of prescription stimulants in people without ADHD (Smith & Farah, 2011; Chamberlain et al., 2010). Yet, the use of stimulant medication among healthy people is on the rise (Smith & Farah, 2011). Thus, it remains an open question what drives the enhancement uses of medications like Adderall and Ritalin. This paper will focus on three non-mutually exclusive candidate explanations of the surprisingly wide-spread enhancement stimulant use given the limited empirical evidence for the efficacy of cognitive enhancement: the possibility that use is related to users' attention problems, low motivation, or suboptimal study habits. Our goal was to examine if users and non-users differ on these dimensions – a first step towards investigating, down the road, important directional causal questions: Do users self-medicate undiagnosed attention difficulties? Do they intervene in perceived attention problems despite objectively normal attention? Do they compensate for low motivation or inefficient approaches to learning by resorting to unprescribed stimulants?

Attention Problems and Unprescribed Stimulant Use

Several researchers have suspected attention problems among non-medical stimulant users. This hypothesis has received support from a number of studies of college students, finding higher self-reported inattention and/or impulsivity in users, compared to their non-using peers (Arria et al., 2011; Rabiner et al., 2009; Peterkin et al., 2011; Rabiner et al., 2010). Moreover, longitudinal data have shown that self-

reported attention difficulties in the beginning of college predict prospectively the onset of enhancement use (Rabiner et al., 2010).

Nevertheless, the proposed role of attention problems in non-medical stimulant use is strongly qualified by a research limitation shared by past investigations. Previous studies have solely relied on self-report assessments of attention – measures susceptible to bias (e.g., see Hunt et al., 2011). For instance, users might consciously or unconsciously exaggerate their symptoms to justify self-medication. Alternatively, students surrounded by high achieving peers might perceive their normal attention abilities as deficient. Thus, without converging evidence from self-report and objective neuropsychological testing, it is difficult to infer and explain users' attentional impairment.

The most widely used objective test of attention is the Test of Variables of Attention (TOVA). The TOVA is a continuous performance test, which presents subjects with a sequence of simple geometric figures signaling either a "go" or a "no-go" response. Several strengths of this instrument make it suitable for the objective assessment of attention. Age- and gender-normed standard scores are automatically generated, allowing an inference about the clinical significance of participants' performance. Malingering is detectable through an index of symptom exaggeration, considered positive if relevant conditions are met (e.g., if post-commission responses are quicker than the mean reaction time). The TOVA has better sensitivity and specificity than standard continuous performance tests: Its 22-minute duration prevents above-threshold performance purely due to a compensatory strategy when actual attention difficulties are present. Additionally, the test's non-verbal stimuli help differentiate attention problems from reading disorder (Forbes, 1998; Hunt et al., 2011).

Thus, this test is a suitable instrument to evaluate whether enhancement stimulant users have *objectively* lower attention performance, given their *subjectively perceived or reported* attentional difficulties.

Motivation and Unprescribed Stimulant Use

Aside from optimal attention, non-medical stimulant users might be seeking increased motivation to study. Several lines of research have converged to suggest that stimulants are beneficial for improving motivation. Motivation encompasses a variety of facets, including, but not limited to liking (e.g., enjoying a task) and wanting (e.g., ascribing value to the task outcome; expending effort in a task). Animal research shows that stimulants increase activity in the mesolimbic dopamine system, which is central to motivation (Butcher et al., 1988, Drevets et al., 2001; Volkow et al., 2004). Double-blind, placebo controlled laboratory experiments of stimulant effects in humans have documented elevated self-reported interest in a mathematical task, correlated with change in striatal dopamine (Volkow et al., 2004); increased enjoyment of viewing IAPS images (Wardle et al., 2012); increased expenditure of effort for reward in a laboratory task (Wardle et al., 2011); and a stimulant-related increase in self-reported energy (e.g., de Wit et al., 2000, Costa et al., 2012). A survey from our lab indicated that enhancement users rate stimulants' motivational effects as at least as pronounced as the cognitive ones (Ilieva & Farah, 2013). Thus, a number of experiments, using self-report, behavioral and neural measures, have supported the effects of ADHD medications on motivation in non-clinical samples.

Research on enhancement users' experiences has found that that stimulants' motivational properties are highly sought for. A recent study, based on semi-structured interviews and qualitative analyses, showed that users particularly value the stimulant-

related increases in drive and task enjoyment (Vrecko, 2013). As representative participants noted, "[on Adderall] I didn't want to stop what I was doing until it was completed up to a certain level of my satisfaction," and "You're interested in what you're doing even if it's boring." Structured surveys asking participants to choose among candidate motives for unprescribed stimulant use have found that a majority of users seeks a stimulant-driven increase of energy and task enjoyment (e.g., DeSantis et al., 2008, Bavarian et al., 2013, Teter et al., 2005).

Given their interest in stimulants' motivational properties, might users have overall lower motivation for cognitive tasks than controls? To address this question, we examined users' and controls' subjective experience of the TOVA, focusing on how boring they found the task and how driven they were to do well. Our self-report measure is useful in distinguishing the subjective experience of motivation from attentional performance during cognitive testing. It is also appropriate for a preliminary investigation, which, if yielding significant findings, can substantiate a more comprehensive assessment of more facets of motivation in future.

Study Habits and Unprescribed Stimulant Use

Whether or not they have an attentional disorder or low motivation for their schoolwork, stimulant enhancement users may also seek medication to compensate for poor study habits. We use the term study habits to describe study practices that either facilitate or impede successful and efficient learning. Here, we are interested in study habits at a behavioral level, without attempting to parse out the relative causal contributions of psychopathology, lack of proper instruction and training, low achievement motivation, low self-control or unfavorable situational factors.

Several lines of research converge to suggest the possibility of suboptimal study habits among non-medical stimulant users. Previous work has indicated that users spend less time studying and skip more classes than their non-using peers (Arria et al., 2011, 2013). Cramming for exams and improving study skills have been identified as common motives for unprescribed ADHD medication use (de Santis et al., 2008, Hildt et al., 2014, Peterkin et al., 2011). An inverse relationship has been documented between trait Conscientiousness and unprescribed stimulant use (Benotsch et al., 2013). Taken together, these data raise the possibility that use is associated with the quality of students' study habits – a construct more specific to academic behavior than trait conscientiousness, but, as shown below, more comprehensive than the isolated student behaviors examined previously.

Previous research has identified a number of study practices beneficial for learning. Spaced practice of to-be-learned material leads to longer-term retention than massed practice. Retrieval practice improves memory relative to no practice or to repeated exposure to the same material. Critical analysis of the studied material (e.g., interpreting and interconnecting information) is another strategy shown to benefit retrieval (see Roediger & Pyc, 2012; Dunlosky et al., 2013; Bjork, Dunlosky & Kornell, 2013, for reviews of the solid body of research that supports the effectiveness of these approaches). Other activities found to correlate with successful learning outcomes in school and at work include: persistence despite failure or boredom, time management, the tendency to work in distraction-free environments, as well as planning and monitoring one's goal-directed behavior (Sitzman & Ely, 2011; Crede & Phillips, 2011). The small to moderate size of the correlations with learning outcomes does not necessarily discount the importance of these study practices: they may be an important

determinant of success, though acting only in conjunction with intelligence and other factors and effective only if applied properly (Bjork et al., 2013).

We asked whether users and controls differ on this broader, more comprehensive array of study habits. To address this question, we compiled our own set of self-report items borrowed from several existing scales (see Appendix), with the aim of assessing: 1) study habits previously shown to effectively promote learning and achievement; and 2) study habits (e.g., note-taking and class participation) that appeared important for academic success to three independent research staff members who reviewed the published scales. Despite the availability of a multitude of published measures of study habits in the literature (e.g., Pintrich et al., 1991; Schmeck et al., 1991; Nonis & Hudson, 2011; Kornell & Bjork, 2007; Grendler & Garavalia, 2000; Christopolous et al., 1987; Biggs, 1987, etc.), we decided against directly using one of these measures, because none met fully our first and main aim, as described above.

The Present Study

The goals of the present study were to examine attention, motivation and study habits in stimulant enhancement users, relative to controls with no history of ADHD medication use. We conducted a multimodal assessment of attention, combining a subjective measure with an objective neuropsychological test. We predicted lower self-reported attention among users, while making two alternative hypotheses about user-control differences on objectively measured attention. If use relates to true attention problems, we expected to see lower TOVA performance in users than controls. Alternatively, if use is more strongly driven by perceived functioning than by objective problems, we expected an interaction pattern, indicating relatively lower functioning on self-report in users than controls, despite comparable objective performance in the two

groups. We further predicted lower level of self-reported motivation among users for the duration of the TOVA, as well as less optimal self-reported study habits, relative to controls. We were interested both whether these outcomes distinguish users from controls, as well as whether they remain significant even after holding constant previously documented group differences on depression, anxiety and substance use (Dussault & Weyandt, 2011; Weyandt et al., 2009; Rabiner et al., 2010; Teter et al., 2010; Arria et al., 2013; McCabe et al., 2004; McCabe & West, 2013). Examining the functioning of enhancement users is a necessary first step towards asking, down the road, what factors contribute causally to non-medical ADHD medication use.

Method

Participants

The analyzed data is from 128 participants, a sample size selected to attain 80% power of detecting medium-sized effect (Cohen's *d* >= 0.5) in our main analyses. The sample consisted of 61 enhancement users of prescription stimulants (27 female, 34 male) and 67 controls (37 female, 30 male), who reported no lifetime prescription stimulant use. All participants were young adults (age range 18-28, *M* = 20.95, *SD* = 2.05) who denied a history of ADHD diagnosis. Participants were recruited through university-affiliated recruitment web-sites and flyers on university campuses in Philadelphia. The project was advertised as "a research study comparing users of unprescribed ADHD medication to people who have never used such drugs."

In addition to this final sample of 128 participants, 48 more participants began the study without completing it or without being included in the analyses. Of these 38, 24 participants (14 users and 10 controls) dropped out after completing part of the

study[3]. Additional 24 participants were excluded for the following reasons: possible symptom exaggeration on the TOVA (n = 4); inconsistent information about history of ADHD at different assessment points (n = 1); inconsistent information about enhancement use (admitted vs. denied) at different assessments (n = 5); five or more alcoholic drinks the evening before the TOVA (n = 7); four or fewer hours of sleep the night before the TOVA (n = 3); the equivalent of a cup of coffee or more before the TOVA, given no typical caffeine intake[4] (n = 3); runs of sequential omission errors (a rare pattern of performance typical of narcolepsy and seizure disorders, n = 1). The latter criterion was applied because we were interested in generalizing our finding to a relatively typical population of young people. Participants who took medications with stimulant properties (e.g., stimulant medications, modafinil, atomoxetine, bupropion) before the TOVA were ineligible, but none presented to the lab meeting this criterion.

Procedure

The study began with a screening survey, excluding people with an ADHD diagnosis, as well as people outside of the 18-30 age range. Potential participants were also asked about lifetime use of prescription stimulants (yes vs. no). Depending on user status, they were directed to two separate sign-up lists. This early distinction between users and non-users allowed us to keep the number of enrolled participants roughly equal between groups.

[3] In the sample of 24 people who dropped out, users reported higher levels of depression ($p = 0.04$), anxiety ($0.03 > ps > 0.04$), substance use ($p < 0.01$), and attention problems ($0.11 > p > 0.08$) than controls. Users also reported less optimal study habits ($p = 0.18$) than controls. As will be shown below, these trends suggest that user-control relationships in noncompleters, at least based on these available data, are consistent with our findings among completers.

[4] TOVA performance has been documented to be sensitive to caffeine intake only among those who do not habitually take caffeine (Hunt et al., 2011)

The initial phase of the actual study consisted of an online battery of self-report assessments on study habits, attention, anxiety, depression, and substance use, administered in that order. A separate second session began with the Test of Variables of Attention (TOVA), continued with participants' self-report on their motivation during the computerized test and a self-report on the incidence of their enhancement stimulant use. The session concluded with a report on medication use, caffeine intake, alcohol and illicit substance intake, as well amount of sleep before testing and history of ADHD diagnosis. A day prior to the study, subjects had been contacted with instructions to take their usual amount of sleep before testing and to refrain from taking more caffeine than usual on the test day. 74 participants (36 users) were tested in lab by blind experimenters; for the remaining 54 participants (25 users) testers were not blind to user status[5].

Materials

Main measures

Enhancement Stimulant Use. Participants indicated the number of occasions of unprescribed ADHD medication use in the past month, past year and in their lifetime. The measure, adapted from Teter et al. (2010), read as follows: "On how many occasions have you used **ADHD medication** (e.g., Adderall, Ritalin, or other), **without a prescription**, to help you **do well at school** and/or work?" Our main analyses were based on the incidence of lifetime use (given its greatest range among the three

[5] To assess possible experimenter effects on each of the TOVA indexes, we examined the interactions between user status and tester blindness, based on a series of two-way independent-samples ANOVAs. No significant interactions emerged. There were no effects of tester blindness on any of the TOVA variables within the separate subsamples of users and controls, according to the results of t-tests and Mann-Whitney tests (all ps > 0.38).

measures). Sensitivity analyses using data on past-year and past-month use were also conducted.

***Barkley & Murphy ADHD Current Symptom Scale*.** This self-report ADHD assessment instrument incorporates scales of inattention (9 items) and impulsivity (9 items), as well as an evaluation of symptom-related impairment. The scale items correspond to DSM-IV-defined ADHD symptoms, with wording adapted for adult populations. Participants indicate the frequency of each symptom on a 0 ("Never or rarely") – 4 ("Very often") scale. An indication of frequent or very frequent manifestation of at least 6 inattention *or* 6 hyperactivity symptoms meets the scale's cutoff for clinically significant impairment. The scale has demonstrated excellent positive predictive value (0.8-1) but limited negative predictive value (0.3) in previous research (O'Donnell et al. 2001). Thus, diagnosis cannot be established purely based on the results of the scale, in the absence of report from other informants on the nature, severity, pervasiveness and childhood onset of the diffculties (Murphy & Adler, 2004).

Test of Variables of Attention (TOVA). The TOVA is a 21.6-minute continuous performance test. Participants are presented with a sequence of briefly flashed simple geometric figures, requiring participants either to press a button as quickly as possible or to withhold responding. The first half of the test taxes inattention, given infrequent target presentation, based on a target:non-target ratio of 1:3.5. The second half taxes impulsivity, given frequent target presentation, based on the reverse target:non-target ratio of 3.5:1. Throughout the test, stimulus presentation is 100 ms and interstimulus interval is 2s. The TOVA provides a symptom exaggeration index, which is considered positive if at least two of the following four criteria are met: quick post-commission responses, slow commission errors, extreme omission, commission or reaction time

variability scores. The TOVA's specificity and sensitivity in identifying ADHD have been estimated to range between 67%-86% (Greenberg & Waldman, 1993; Forbes, 1998; Schatz et al., 2001).

Our dependent variables included three measures of inattention: omission errors, reaction time variability, and reaction time; one measure of impulsivity: commission errors; and the overall attention performance index (API) score. The API reflects a linear combination of reaction time in the first half of the test, sensitivity (d') in the second half of the test, and reaction time variability over the duration of the total test. This is a combination of variables, previously indicated to best predict ADHD (Greenberg & Waldman, 1993). The API falls on a -10 to +10 range, where negative numbers are suggestive of clinically significant attention problems. The remaining dependent measures are automatically reported as standard scores, with higher standard scores indicating better performance.

Motivation and subjective experience during the TOVA. Participants rated their experience of completing the TOVA test on six scales. Four of these items assessed two aspects of motivation: boredom ("unpleasant" – "enjoyable," "very fun" – "very boring,") and drive ("not motivated to do well" – "very motivated," effort invested in the task: "as much as possible" – "none at all"). Two items assessed how difficult and how tiring participants found the test ("easy" – "difficult," "very-exhausting" – "not tiring at all"). All items were scored on 5-point scales. These measures were completed twice: once at the end of the short TOVA practice test, and once at the end of the full actual TOVA test.

***Study Habits*.** A 34-item self-report measure assessed a variety of study habits, including self-testing and rehearsal, spaced practice, effort and persistence, critical

analysis of the material, time management, preference for work-appropriate spaces, self-monitoring of goal-directed activities, class attendance, assignment completion, time spent studying, among others. Participants were presented with statements, each describing a study habit, and asked to indicate how frequently they rely on that study habit, using a 0 (Never) – 4 (Always) scale. Items were compiled from previously published scales on study habits. In our sample, the scale had good-to-excellent internal consistency: Cronbach's alpha = 0.88. Furthermore, in this sample, the measure of study habits was significantly associated with GPA (r = 0.38, ps < 0.01), depression (BDI: r = -0.31, Spearman's rho = -0.23, ps =< 0.01), trait anxiety (STAI-general: r = -0.30, Spearman's rho = 0.30, p < 0.01), and self-reported attention (Current Symptom Scale Total Score: r = -0.35, Spearman's rho = - 0.29, ps < 0.01; Current Symptom Scale - Inattention Subscale: r = -0.47, Spearman's rho = 0.40, ps < 0.01; B Current Symptom Scale - Impulsivity Subscale: r = -0.18, Spearman's rho = 0.04; rho = - 0.17, p = 0.06). We found no correlations between Study Habits and any of the TOVA indexes.

Secondary Measures

Secondary measures reflected demographics, as well as several control variables (e.g., depression, anxiety, substance use).

Beck Depression Inventory – II (BDI). The BDI is a measure of depression severity, tailored to reflect the DSM-IV diagnostic criteria. Each of the 21 irems on the BDI is rated on a 0-3 severity scale for a maximum score of 63. Conventionally, scores in the ranges 0-9; 10-19; 20-29 and 30-63 reflect, respectively, minimal, mild, moderate and severe depression. The BDI has excellent reliability and validity (e.g., Steer et al., 1999; Storch at al., 2004).

State Trait Anxiety Inventory (STAI). The STAI is a widely used self-report assessment of anxiety, from which we selectively focused on the 20-item subscale reflecting trait anxiety. Participants were asked to rate the extent to which they experience various anxiety symptoms (e.g., nervousness, insecurity) on a 0 (not at all) – 3 (very much so) scale. The test has high test-retest reliability and correlates highly with other anxiety questionnaires (Spielberger, 1983), although it does not consistently differentiate anxiety from depression (Bados et al., 2010, Balsamo et al., 2013).

Substance Use. To assess substance use, participants were given a list of addictive, commonly abused substances, some of which were also identified with a street name. These included: tobacco, marijuana, MDMA ("molly" or "ecstasy"), cocaine, hallucinogenic mushrooms, LSD, heroin, methamphetamine, opioids, unprescribed opioid painkillers, PCP ("angel dust"), hashish, unprescribed barbiturates or benzodiazepines, and inhalants. For each substance, participants indicated the number of occasions of use in their lifetime.

Other demographic and control variables. Data were also collected on participants' gender, undergraduate institution, GPA, and current occupation. To examine some situational factors, potentially affecting TOVA performance, we administered a list of open-ended questions about medication intake (type and dose) within 24 hours before the TOVA; caffeine intake (type and amount of caffeinated drink) on the day of the TOVA, as well as on a typical day; and alcohol and substance use (type and amount of substance) within 24 hours before the TOVA. We also inquired about the number of hours participants slept the night before the TOVA. Before the objective attention test, all participants confirmed that their vision was normal or corrected-to-normal.

Results

Data Distributions and Choice of Parametric vs. Non-parametric Tests

Several of our main variables of interest had non-normal distributions, as indicated by a series of significant Shapiro-Wilks tests. Non-normally distributed variables included all indexes of objective attention (TOVA: omissions, commissions, reaction time variability, reaction time and API) and subjective attention (Current Symptom Scale: inattention subscale, impulsivity subscale and total score), as well as the BDI, STAI-general, our measures of substance use incidence and amount of sleep pre-TOVA. These distributions were skewed, in some cases pronouncedly so: a majority of data indicated uniformly high functioning, while increasingly fewer subjects showed (or reported) increasingly greater problem severity. We attempted several transformations (square root, square, ln and log10) of the raw or the reversed scores, but without attaining normality. Hence, our main analyses relied on non-parametric tests. Secondarily, we conducted parametric procedures with untransformed data. Measures of study habits and motivation were normally distributed, allowing analyses using parametric procedures only.

Handling of Outliers

We winsorized all data by substituting the three highest and three lowest data points (4.7% of the data) with the next most extreme data point.

Handling of External Variables

Based on consistent past findings (Dussault & Weyandt, 2011; Weyandt et al., 2009; Rabiner et al., 2010; Teter et al., 2010; Arria et al., 2013; McCabe et al., 2004; McCabe & West, 2013), we assumed that elevated levels of depression, anxiety, and

substance use are characteristic of users. Thus in our main analyses, we do not statistically control for these variables, in order not to partial meaningful group variance out of the analyses. If we held constant the values of depression, anxiety and substance use between groups, we run the risk of obtaining findings unrepresentative of a substantial proportion of users (Miller & Chapman, 2001). However, in a secondary set of analyses, we do enter the third variables as predictors in the model, to assess if users and controls differ on attention, motivation and study habits above and beyond their previously documented differences on depression, anxiety, and substance use.

Subsets of Data Analyzed

Our main analyses, which are reported below, were conducted based on the full sample of eligible participants. In addition, we replicated these analyses in two subsets of participants. First, we excluded participants (n = 5) who disclosed having used enhancement medication only once in their lifetime. Our reasoning was that one-time users might be unrepresentative of people who use continually, for instance, by finding stimulants unhelpful, by experiencing side effects as intolerable, or by functioning relatively more highly than other users in the areas of interest in the present study. Secondly, we replicated our analyses in users with API scores within normal limits (n = 109). User-control differences on motivation and study habits were most likely to be detected in this subsample, as was the evidence for disparities between perceived attention problems and high functioning on objectively measured attention. The results of these two sets of secondary analyses are only reported when different in direction or significance level from the findings of the main analyses.

Characteristics of Enhancement Users

A chi square test for independence indicated a non-significant relationship between gender and user status ($\chi^2 = 1.53$, $p = 0.22$). Contrary to intuitive expectations, there was a borderline significant trend for users to report more time having slept the night before testing ($M = 7.40$, $SD = 1.19$) than non-users ($M = 7.02$, $SD = 0.97$): t (122) = 1.93, $p = 0.06$. Positive correlations between hours of sleep and lifetime incidence of use were significant ($r = 0.18$, p = 0.05; Spearman's rho = 0.21, $p = 0.02$). Users were more likely than controls to have taken a cup of coffee or more (or a roughly equivalent amount of another caffeinated drink) on the test day: χ^2 (1) = 5.29, $p = 0.02$. Consistent with past findings (Dussault & Weyandt, 2011; Weyandt et al., 2009; Rabiner et al., 2010; Teter et al., 2010; Arria et al., 2013; McCabe et al., 2004; McCabe & West, 2013), users reported higher levels of depression (t (126) = 3.28, $p < 0.01$; Mann-Whitney U = 2,633.50, z = 2.82, $p < 0.01$), trait anxiety (t (127) = 3.35, $p < 0.01$; Mann-Whitney U = 2,677.50, z = 3.03, $p < 0.01$), and substance use (t (127) = 5.75, $p < 0.01$, Mann-Whitney U = 3,141.50, z = 5.28, $p < 0.01$) than controls. Finally, self-reported GPA was lower among enhancement users ($M = 3.29$, $SD = 0.38$) than controls ($M = 3.55$, $SD = 0.41$), t (124) = 3.67, $p < 0.01$. Depression, anxiety, substance use and GPA were also significantly correlated with lifetime use, based on both parametric and non-parametric tests.

Stimulant Enhancement Use and Attention

Non-parametric and parametric procedures consistently showed higher level of *self-reported* inattention (U = 2,619.50, z = 2.76, $p < 0.01$, one-tailed, $d = 0.50$; t (126) = 2.95, p < 0.01, one-tailed, $d = 0.52$) and impulsivity (U = 2,637, z = 2.99, $p < 0.01$, one-tailed, $d = 0.54$; t (126) = 3.26, $p < 0.01$, one-tailed, $d = 0.57$) on the Current Symptom

Scale among users than controls. Accordingly, users had higher total scores on this self-report scale: U = 2,669.50, z = 2.99, $p < 0.01$, one-tailed, $d = 0.54$; t (126) = 3.36, $p < 0.01$, one-tailed, $d = 0.59$. Lifetime enhancement use correlated with subjectively perceived attention problems: $r = 0.31$, Spearman's rho = 0.28, ps < 0.01, one-tailed, for the Inattention subscale of the Current Symptom Scale; $r = 0.35$, Spearman's rho = 0.29, ps < 0.01, one-tailed, for the Impulsivity subscale of the Current Symptom Scale; $r = 0.35$, Spearman's rho = 0.31, ps < 0.01, one-tailed, for the total score of this scale. Qualitatively similar patterns emerged when correlating the measures of attention with past-year and past-month enhancement use.

On the *objective* test of attention, independent-samples Mann-Whitney tests showed a higher number of omission errors (the measure with the most pronouncedly skewed distribution) among users than controls (U = 1,610, z = 2.09, $p = 0.02$, one-tailed, $d = 0.38$), as well as lower overall attention performance index on the TOVA among users (U = 1,693, z = 1.67, $p = 0.05$, one-tailed, $d = 0.30$). These differences emerged, even though, as shown above, users had slept slightly longer the night before testing and were more likely to have taken the equivalent of a cup of coffee before testing. In contrast, when a series of between-subjects one-way ANOVAs were applied to the skewed data distributions, no significant differences emerged between the groups on any index of objective attention (irrespective of whether we controlled for sleep and caffeine before testing).

We also examined the correlations between incidence of enhancement use, on the one hand, and each index of objective attention. According to the results of non-parametric tests, omission errors were weakly correlated with lifetime enhancement use (Spearman's rho = 0.15, $p = 0.05$, one-tailed), and past-year use (Spearman's rho =

0.19, p = 0.02, one-tailed). The relationship with past-month use did not reach significance, possibly due to the restricted range of this measure. No other correlations between enhancement use and the remaining indexes of objective attention emerged significant, based on either non-parametric or parametric tests (all ps > 0.08, one-tailed) and irrespective of controlling for sleep and caffeine before the TOVA in corresponding parametric regression analyses. Thus, unprescribed stimulant use was associated with objectively measured inattention, but less strongly and consistently than with perceived attention difficulties.

Are the discrepancies between users' and controls' attention significantly more pronounced on subjective than on objective tests? We conducted a series of three mixed-model ANOVAs with test type (subjective vs. objective) as a repeated-measures factor and user status (users vs. controls) as a between-subjects factor. Dependent measures in each of these three analyses were the following pairs of indexes: 1) API (TOVA) and total score of the Current Symptom Scale; 2) a linear composite of Omissions plus RT Variability (TOVA) and Inattention subscale (Current Symptom Scale); 3) Commissions (TOVA) and Impulsivity (Current Symptom Scale). To convert the objective and subjective data to a common scale, we converted all outcomes to z-scores with consistent directionality. We found a significant interaction on the tests of impulsivity, such that objective scores were very similar between users and controls, while users described themselves as more impaired than controls on self-report: F (1, 126) = 4.48, p = 0.04. The same trends emerged on tests of inattention and of overall attention performance, but the interactions did not reach significance (0.15 < all ps < 0.101). However, when participants who had used unprescribed stimulants only once were excluded, the interactions between user status and attention test type emerged significant on both impulsivity and inattention, showing comparable performance on the

objective test between the two groups, but lower perceived attention among users than controls (for inattention subtests: $F_{interaction}$ (1, 121) = 4.78, p = 0.03; for impulsivity subtests: $F_{interaction}$ (1,121) = 7.91, p < 0.01; for overall attention performance: $F_{interaction}$ (1, 121) = 3.83, p = 0.053, see Fig. 3.1.).

We further asked whether stimulant enhancement use is disproportionately more common among people whose scores fall in the attention tests' range of clinically significant impairment. 9 users and 2 controls scored in the clinical range of the Current Symptom Scale. A significant chi square test for independence showed that users are significantly more likely to have above-threshold self-reported attention difficulties: chi χ^2 = 5.63, p = 0.02. In contrast, we found that 9 users and 10 controls performed below the API's clinical cutoff on the TOVA (χ^2< 0.01, p = 0.98). In sum, enhancement use appeared disproportionately common among people with self-reported attentional problems. Those with objectively measured attention problems were equally likely to report and deny stimulant enhancement use.

Self-reports on participants' experience of the TOVA showed positive correlations between the tendency to describe the test as difficult, on the one hand, and past-year (r = 0.19, p = 0.02, one-tailed; Spearman's rho = 0.19, p = 0.02, one-tailed) and past-month use (r = 0.17 , p = 0.03, one-tailed; Spearman's rho = 0.18, p = 0.03, one-tailed). No significant correlations with lifetime use emerged. In addition, no correlation emerged between incidence of use (lifetime, past-year and past-month) and participants' tendency to describe the attention test as tiring.

Stimulant Enhancement Use and Motivation

An independent-samples t-test indicated that users reported lower overall motivation during the TOVA test than controls: t (126) = 3.09, $p < 0.01$, one-tailed, d = 0.54, based on a composite of the four motivation-related items. A closer look at specific sub-groups of items indicated that users described the test as more boring (t (126) = 2.83, $p < 0.01$, one-tailed, d = 0.50, based on a composite of the items "very fun" – "very boring" and "unpleasant" – "enjoyable") and, reportedly, were less driven to do well (t (126) = 2.11, p = 0.02, one-tailed, d = 0.37, based on a composite of the items "not motivated to do well" – "very motivated" and effort expended on the task: "as much as possible" -- "none at all"). Accordingly, the incidence of lifetime stimulant enhancement use was inversely correlated with the test motivation composite (r = -0.26, Spearman's rho = 0.28, ps < 0.01, one-tailed), task enjoyment (r = -0.24, rho = 0.28, ps < 0.01, one-tailed) and drive (r = -0.17, p = 0.03, one-tailed; Spearman's rho = -0.18, p = 0.03, one-tailed). Correlations between these motivation indexes and past-year use replicated the reported findings, whereas correlations with past-month use failed to reach significance, possibly due to the more restricted range of this measure.

Another way of examining users' motivation during the TOVA entails asking if their motivation ratings' linear composite decreased more dramatically over the duration of the TOVA test, relative to the control group. We conducted a mixed-model ANOVA with user status as a between-subjects factor and motivation assessment time point (after TOVA practice; after the full test) as a within-subjects factor. No significant interaction emerged either based on the full sample, or after one-time users were excluded (ps $_{\text{interaction}}$ > 0.19). However, when only analyzing data from people with API scores within normal limits, the interaction between user status and pre-post

assessment emerged significant (F (1,105) = 4.21, p = 0.04). This interaction revealed that the drop in motivation from the beginning to the end of the TOVA was greater among users than controls with normal attention functioning.

Stimulant Enhancement Use and Study Habits

An independent-samples t-test indicated that users reported less optimal study habits than controls: t (126) = 2.65, p < 0.01, one-tailed, d = 0.48. Accordingly, ratings on study habits quality correlated inversely with the incidence of lifetime stimulant enhancement use: r = -0.20, p = 0.01, one-tailed, and past-year enhancement use: r = -0.23, p < 0.01, one-tailed. No significant correlation with past-month use emerged (r = -0.13, p = 0.07, one-tailed), possibly due to the relatively more restricted range of this scale.

Analyses Controlling for External Variables

In addition to describing the relationships between unprescribed stimulant use, on the one hand, and attention, motivation and study habits, on the other hand, we asked if these associations are significant after holding constant factors which have been previously documented to differ between users and controls: depression, anxiety and substance use. In other words, we were interested if use is associated with attention, motivation and study habits over and above what could be accounted for by depression, anxiety and substance use. We conducted a series of between-subjects ANOVAs for each outcome, with the control variables (trichotomized to circumvent distribution skewness) and user status as between-subjects factors. For analyses with TOVA indexes as outcomes, we additionally entered caffeine use (dichotomous) and sleep before the TOVA (trichotomized) as independent variables. In these analyses, we

found trend-level relationships between user status, on the one hand, and subjective and objective attention and study habits, on the other ($0.16 < p < 0.07$, one-tailed). The motivation composite remained related to user status after depression, anxiety and substance use were held constant: F (1, 127) = 4.96, p = 0.01, one-tailed (for the main effect of user status).

Discussion

Our study examined the psychological profile of people without diagnosed ADHD who use unprescribed stimulants to enhance their academic or professional performance. Aside from replicating previous findings of lower self-reported attention among users relative to their non-stimulant using peers, we extended these data in important ways. Specifically, we obtained evidence for somewhat lower functioning in users than controls on an objective, neuropsychological measure of attention, a measure not susceptible to the biases inherent in self-report. Furthermore, we found that, relative to controls, users describe their study habits as poorer and report lower motivation during laboratory attention testing. The motivational differences between the two groups remained significant even when statistically controlling for depression, anxiety and substance use.

There are several ways to interpret the finding that the relationship of enhancement use with attention and study habits emerged only at trend level when holding constant depression, anxiety and/or substance use. This finding may indicate an absence of user-control differences on attention and study habits *beyond* what is already captured by depression, anxiety and substance use. Alternatively, attention and study habits may be related to use above and beyond the controlled variables, but these relationships might be weak and detectible only in larger samples. (In the

abstract, there is also the possibility that attentional difficulties and poor study habits merely co-occur with what actually predicts use, without being causally related to use. This possibility is unlikely, given previous findings showing that users identify the optimization of attention as the primarily motive for self-medication with stimulants, Boyd et al., 2006; Hall et al., 2005; Teter et al., 2005, 2006, and study habits as another, though less frequently endorsed, motive, Rabiner et al., 2009). Regardless of their relationship to the variables we controlled for, attention and study habits do appear more compromised on average in users than controls.

Although both perceived and objectively measured attention difficulties were associated with unprescribed ADHD medication use, use appeared more strongly and robustly related to subjectively perceived attention functioning. Differences between users' and controls' attention were larger in size on subjective than on objective measures. Whereas consistently associated with all self-report attention measures, use remained unrelated to some TOVA indexes (incl. commission errors, the primary measure of impulsivity, and RT variability, one of the indicators of inattention). While individuals with clinically significant attention difficulties on self-report were disproportionately likely to use enhancement stimulants, participants with clinically significant difficulties on objectively measured attention were equally likely to report and deny use. Taken together, these patterns indicate stronger relationship of unprescribed ADHD medication use with a perception of attentional problems than with actual difficulties.

What might explain the relatively weak relationship of objective attention to non-medical stimulant use? Although some users may seek to medicate clinically significant attention problems or optimize normal attention, a majority may be relying on stimulants

to medicate poor study habits, demotivation, anergia and fatigue, possibly in some cases secondary to clinical or subclinical mood, anxiety or substance-related problems. If primarily emerging in these contexts, users' attention difficulties would be expected to be relatively mild (i.e., not as pronounced as one would expect in ADHD), and the differences in users' and controls' functioning would be relatively subtle.

While likely motivating the use of enhancement stimulants, low levels of perceived attention, task motivation and study habits may also, in part, be a product of unprescribed stimulant use. For instance, some users might be justifying self-medication by perceiving or reporting attention difficulties; some might be deducing attentional impairment from the fact that the medication feels effective. At the same time, the availability of Adderall as a study aid might be reducing the perceived need for maintaining self-regulated study habits. Given our study's cross-sectional design, a third variable may be explaining some of the documented relationships: people who admit to illicit medication use may be less prone to socially desirable responding than the rest of the sample. This could accounting in part for users' less favorable self-reports. The present study cannot distinguish between these explanations, but, as outlined below, paves the way for future longitudinal and intervention studies, which can establish the direction and causality of the examined relationships.

Limitations

Several limitations of the present study require consideration. First, we relied on self-report to establish the incidence of unprescribed stimulant use and the absence of previous ADHD diagnosis. One could imagine that, motivated to get into the study, participants might have concealed ADHD diagnosis or dishonestly indicated a history of unprescribed medication use. Conversely, to avoid academic or legal repercussions, or

driven by social desirability, some might have concealed illicit prescription stimulant use. In the low-likelihood case of non-negligible *systematic* bias for dishonest reporting, our findings might be describing inaccurately the correlates of use. In the absence of systematic bias in reporting, the signal-to-noise ratio in our data might be suboptimal.

Second, 24 participants dropped out of our study (i.e., completed the online survey but did not return for the TOVA), raising a question whether users with the most impaired objective attention remained untested. Fortunately, the number of non-completers was roughly comparable between the user and control group (14 users vs. 10 controls), reducing (though not eliminating) the possibility of systematic between-group differences in the functioning of missed cases.

Finally, given that the majority of our sample consisted of students (124 out of 128) completing or having completed their undergraduate degree at the University of Pennsylvania (114 out of 128 participants), the generalizeability of our finding to other occupations or other undergraduate institutions is an open question.

Future Directions

The present study opens up important avenues for future research. First, subsequent investigations can examine the relationship of unprescribed stimulant use to complementing measures of motivation and study habits. For a behavioral assessment of specific learning strategies, one could employ (or modify) Son & Kornell's (2009) paradigm, which asks participants to study word pairs for a subsequent test and observes their learning strategies (e.g., spaced vs. massed practice) in the lab. Modifications of this procedure could evaluate the previously unexamined relationship of enhancement use to individual study habits, including time allocation for task-

oriented activity and choice of self-testing (vs. passive review of the material). Analogously, subsequent research can assess the relationship of enhancement use with a comprehensive array of motivation-related functions. Motivation encompasses a number of facets, measurable through self-report and/or behavioral tests. Examples include the tendency to expend effort for reward (as measured through the behavioral EEfRT task, Treadway et al., 2011); trait drivenness (assessed by the Drive subscale of the Behavioral Activation scale) and positive expectancy for one's performance. Thus, the present study substantiates a more comprehensive assessment of the relationship of stimulant self-medication to motivation and study habits, using measures of various modalities.

Additionally, future investigations may ask questions about the directionality and causality of the relationships examined here. Longitudinal research can examine whether objective attention, motivation and study habits assessed in late adolescence prospectively predict the onset of non-medical stimulant use in young adulthood. Intervention studies can provide insights into the causal roles of motivation and study habits in non-medical stimulant use, while at the same time illuminating the approaches to reducing this risky behavior. Our study raises the possibility of several potentially effective interventions. Past research suggests that students harbor misconceptions about what study habits are optimal (Kornell & Bjork, 2008). A psychoeducational intervention addressing these misconceptions may improve study ativities and, potentially, reduce non-medical stimulant use. Cognitive-behavioral interventions may also be helpful in enhancing users' motivation and study habits, while reducing academic impairment due to depression and anxiety. Research on the effects and mechanisms of these interventions (e.g., in comparison to a control condition, such as psychoeducation on the risks and uncertain benefits of stimulant self-medication) can

be informative of the causal roles of motivation and study habits in enhancement stimulant use.

Future studies can investigate other aspects of enhancement users' psychological profile, including possible weaknesses (e.g., planning and problem-solving) in need of intervention and potential strengths to draw from in compensating for these weaknesses. Finally, given our small subsample of people with possible clinically significant attention problems who reported no ADHD diagnosis, future research and policy should identify and intervene into the barriers to appropriate diagnosis and treatment seeking.

Conclusion

Despite its limitations, the present study extends the previous literature on the correlates of stimulant enhancement in important ways. It shows that non-medical stimulant use is more strongly related to a subjective perception of attention difficulties, inefficient study habits and low task motivation than with actual attentional impairment. The present research has important implications for future research into the causal mechanisms of unprescribed ADHD medication use and into the interventions for discouraging this practice.

CHAPTER 4

GENERAL DISCUSSION

The present project examined stimulants' cognitive enhancement effects in healthy people and the psychological profile of non-medical stimulant users. Study 1, an adequately powered double-blind, placebo-controlled experiment, found no enhancing effect of amphetamine on inhibitory control, working memory, episodic memory, convergent creativity, perceptual intelligence, and a standardized achievement test, despite evidence for subjectively perceived enhancement. No moderating effects of baseline performance or COMT genotype were detected. These findings suggest that drug effects on examined functions are either small or null. In Study 2, we conducted a meta-analysis to distinguish between these two possibilities. Our results showed overall small effects of amphetamine and methylphenidate, based on 47 double-blind, placebo-controlled experiments on inhibitory control, working memory and episodic memory. Given the absence of conclusive evidence for practically significant stimulant effects in healthy people, we conducted Study 3 to address three candidate explanations of the increasing popularity of prescription stimulants' non-medical use. Users reported lower motivation during a laboratory cognitive task and described their everyday study habits as poorer than a control group with no history of stimulant use. In addition, non-medical stimulant use was more strongly related to a perception of compromised attention than to deficits in objectively measured attention. Taken together, these data imply that enhancement users struggle with below-average functioning in one or several cognitive, affective and behavioral domains, compensating for these problems with an illicit intervention of uncertain practical significance.

The present research extends the previous literature in important ways. Study 1, unlike the majority of preceding experiments, had sufficient statistical power to detect medium-sized effects. It was also designed to rule out artifactual explanations of previously documented moderation effects. Whereas previous investigations were restricted to laboratory measures, Study 1 also assessed the effects of an enhancement drug on an ecologically valid test (the SAT). Study 2 built on the only previously published meta-analysis of stimulant enhancement effects (Repantis et al., 2010), by: 1) incorporating more studies (47 vs. 17 in the published meta-analysis); 2) examining the effects of the two most commonly used enhancement stimulants (while Repantis et al., 2010 did not study amphetamine effects); 3) conducting tests of moderation; and 4) assessing the evidence for publication bias. Study 3 complemented previously used subjective assessments of enhancement users' attention with a converging neuropsychological test – a tool not free from the biases inherent in self-report. In addition, Study 3 examined previously unaddressed questions about the psychological profile of enhancement users, with a focus on their study habits and task motivation.

These investigations have important practical and ethical implications. Enhancement stimulants do not appear to pronouncedly optimize high-functioning people's capacity for intelligence, memory and executive functions. It remains an open question whether these medications' effects in healthy people are practical significant under given conditions (e.g., after repeated use), in specific contexts (e.g., task novelty) or in a particular subgroup. Thus, users of unprescribed stimulants appear to pursue uncertain benefits, while exposed to well-established risks, including risks for abuse, dependence, and cardiovascular problems. Stimulants may exacerbate functioning in some people, elevating anxiety in prone individuals, or demotivating persistent, self-

regulated study habits by promising a study aid for a last-minute all-nighter. Costs of non-medical use may be substantial for people self-medicating clinically significant problems with stimulants – a practice which may deter from seeking appropriate medical supervision for a potentially impairing problem.

Thus, aside from above-mentioned outstanding questions about the individual differences moderators and context-specificity of stimulants' effects, the present studies suggest a further avenue for future research. The growing public interest in cognitive enhancement raises the question whether there are interventions that optimize cognition more effectively than unprescribed stimulants. The previous literature has identified a number of candidates, including caffeine, aerobic exercise, meditation, cognitive-behavioral interventions, psychoeducational techniques. Comparing the effects of these enhancement approaches on various facets of cognition and identifying which intervention suits best which group of individuals may be a promising future avenue for the study of cognitive enhancement.

APPENDIX

Study Habits Questionnaire

1. I participate in class discussions even when the instructor does not call on me.
2. When I study for a class I practice saying the material to myself over and over.
3. I ask myself questions to make sure I understand the material I have been studying.
4. I work through practice exercises and sample problems.
5. When working outside of class, I know how to plan my time to get everything done.
6. I don't take all of the notes I should take.
7. When studying outside of class, I keep track of how much time I need to get the work done.
8. I review course material periodically
9. I cram for exams.
10. I spend more time studying than most of my friends
11. I wait till the last minute to complete homework and get ready for exams.
12. I make sure I keep up with the weekly readings and assignments.
13. I attend class regularly.
14. I usually study in a place where I can concentrate on my course work.
15. When reading about research, I like to try out several alternative ways of interpreting the findings
16. I rarely find time to review my notes or readings before an exam.
17. I find it difficult to make much sense of the notes that I take down in class.
18. I often find myself questioning things I hear or read to decide if I find them convincing.

19. When a theory, interpretation or conclusion is presented in class or in the readings, I try to decide if there is good supporting evidence.

20. When I study for a class, I pull together information from different sources, such as lectures, readings, and discussions.

21. I try to understand the course material by making connections between readings and the concepts from the lectures.

22. When I become confused about something I'm reading for a class, I go back and try to figure it out.

23. Before I study new course material thoroughly, I often skim it to see how it is organized.

24. I try to change the way I study in order to fit the course requirements and instructor's teaching style.

25. I try to think through a topic and decide what I am supposed to learn from it, rather than just reading it over when studying.

26. When studying, I try to determine which concepts I don't understand well.

27. When something presented in class is hard to understand, I get everything about it in my notes, so that I could figure it out later.

28. I feel so lazy or bored when I study for my classes that I quit before I finish what I planned to do.

29. I work hard to do well in my classes even if I don't like what we are doing.

30. When course work is difficult, I give up or only study the easy parts.

31. I carefully complete all course assignments.

32. I ask the instructor to clarify concepts I don't understand well.

33. When I can't understand the material, I ask another student in the class for help.

34. I can easily locate particular passages in a textbook when necessary.

TABLES

Table 1.1.

Drug Effect and Interactions, resulting from a series of 2(Drug: MAS; Placebo) x 2 (Drug Order: MAS first; Placebo first) x 2 (Test Version Order: Version 1 first; Version 2 first) mixed-model univariate ANOVAs with repeated measures on the first factor. The dependent variables were scores for 13 measures listed below.

Test (Measure)	Main/Interaction Effects	*df**	*F*	*p*, uncorrected
Face Recognition (number correct)	Drug	1, 40	78	.38
	Drug x Drug Order	1, 40	00	.95
	Drug x Version Order	1, 40	25	.62
	Drug x Drug Order x Version Order	1, 40	6.01	<.01
	Drug Order	1, 40	01	.91
	Version Order	1, 40	.83	.06
	Drug Order x Version Order	1, 40	51	.48
Word Recall (number correct)	Drug	1, 40	10	.76
	Drug x Drug Order	1, 40	.03	.32
	Drug x Version Order	1, 40	10	.76
	Drug x Drug Order x Version Order	1, 40	04	.85

	Drug Order	1, 40	.24	.08
	Version Order	1, 40	28	.60
	Drug Order x Version Order	1, 40	33	.57
Word Recognition (number correct)	Drug	1, 40	82	.37
	Drug x Drug Order	1, 40	.46	.23
	Drug x Version Order	1, 40	.67	.20
	Drug x Drug Order x Version Order	1, 40	.15	.29
	Drug Order	1, 40	.88	.18
	Version Order	1, 40	.15	.02
	Drug Order x Version Order	1, 40	27	.61
Digit Span Backward (number correct)	Drug	1, 38	25	.62
	Drug x Drug Order	1, 38	00	.99
	Drug x Version Order	1, 38	.00	.17
	Drug x Drug Order x Version Order	1, 38	01	.92
	Drug Order	1, 38	18	.67
	Version Order	1, 38	37	.55
	Drug Order x Version Order	1, 38	.92	.06
Digit Span Forward	Drug			.75

(number correct)		1, 38	10	
	Drug x Drug Order	1, 38	.43	.13
	Drug x Version Order	1, 38	.04	.31
	Drug x Drug Order x Version Order	1, 38	29	.59
	Drug Order	1, 38	16	.69
	Version Order	1, 38	50	.49
	Drug Order x Version Order	1, 38	56	.46
	Drug	1, 41	.13	.74
	Drug x Drug Order	1, 41	28	.60
	Drug x Version Order	1, 41	98	.33
Object-2-Back (omissions)	Drug x Drug Order x Version Order	1, 41	11	.74
	Drug Order	1, 41	03	.86
	Version Order	1, 41	00	1.00
	Drug Order x Version Order	1, 41	89	.35
	Drug	1, 38	54	.47
	Drug x Drug Order	1, 38	28	.60
	Drug x Version Order	1, 38	28	.60
Go/No-go (comissions)	Drug x Drug Order x Version Order	1, 38	12	.74

	Drug Order	1, 38	18	.67
	Version Order	1, 38	29	.59
	Drug Order x Version Order	1, 38	34	.56
Flanker (inhibition cost)	Drug	1, 39	20	.66
	Drug x Drug Order	1, 39	13	.72
	Drug x Version Order	1, 39	00	.98
	Drug x Drug Order x Version Order	1, 39	.76	.04
	Drug Order	1, 39	05	.83
	Version Order	1, 39	05	.83
	Drug Order x Version Order	1, 39	17	.68
Remote Associations (number correct)	Drug	1, 42	.56	.12
	Drug x Drug Order	1, 42	01	.94
	Drug x Version Order	1, 42	.10	.16
	Drug x Drug Order x Version Order	1, 42	.44	.01
	Drug Order	1, 42	13	.72
	Version Order	1, 42	14	.72
	Drug Order x Version Order	1, 42	00	.95
Embedded Figures	Drug			.46

(number correct)		1, 32	57	
	Drug x Drug Order	1, 32	01	.93
	Drug x Version Order	1, 32	34	.56
	Drug x Drug Order x Version Order	1, 32	0.68	<.01
	Drug Order	1, 32	.83	.19
	Version Order	1, 32	21	.65
	Drug Order x Version Order	1, 32	88	.35
	Drug	1, 33	01	.91
	Drug x Drug Order	1, 33	05	.80
	Drug x Version Order	1, 33	.92	.18
Raven (number correct)	Drug x Drug Order x Version Order	1, 33	05	.83
	Drug Order	1, 33	01	.93
	Version Order	1, 33	.28	.27
	Drug Order x Version Order	1, 33	.26	.14
	Drug	1, 41	.16	.29
	Drug x Drug Order	1, 41	55	.46
	Drug x Version Order	1, 41	.79	.19
SAT Math (number correct)	Drug x Drug Order x Version Order	1, 41	.56	.04

	Drug Order	1, 41	10	.75
	Version Order	1, 41	.24	.14
	Drug Order x Version Order	1, 41	.02	.32
	Drug	1, 41	47	.49
	Drug x Drug Order	1, 41	.23	.14
	Drug x Version Order	1, 41	37	.55
SAT Verbal (number correct)	Drug x Drug Order x Version Order	1, 41	00	.98
	Drug Order	1, 41	16	.69
	Version Order	1, 41	24	.63
	Drug Order x Version Order	1, 41	.02	.16

******df* **differed between tests due to differences in the number of excluded or missing data points per test.**

Table 1.2.

Means and standard deviations of performance on each dependent measure for the baseline, placebo and mixed amphetamine salts condition.

Task (Measure)	Condition	*N*	*M*	*SD*
Face Recognition (number correct)	Baseline	44	29.05	3.25
	Placebo	44	27.61	4.25
	MAS	44	28.05	4.78
Word Recall (number correct)	Baseline	44	4.25	2.69
	Placebo	44	4.50	4.05
	MAS	44	4.59	3.36
Word Recognition (number correct)	Baseline	44	35.16	4.21
	Placebo	44	34.93	5.65
	MAS	44	34.39	5.04
Digit Span Backward (number correct)	Baseline	42	9.57	2.51
	Placebo	42	10.05	2.70
	MAS	42	10.17	2.80
Digit Span Forward (number correct)	Baseline	42	11.83	1.77
	Placebo	42	12.24	1.59
	MAS	42	12.17	1.67
Object-2-Back (omissions)	Baseline	45	10.38	4.90
	Placebo	45	8.98	4.59
	MAS	45	8.84	5.06

Go/No-go (commissions)	Baseline	42	13.95	5.24
	Placebo	42	15.12	6.20
	MAS	42	14.55	5.50
Flanker (inhibition cost)	Baseline	43	1.16	.05
	Placebo	43	1.16	.06
	MAS	43	1.16	.05
Remote Associations (number correct)	Baseline	46	8.35	2.10
	Placebo	46	7.89	2.50
	MAS	46	8.48	2.18
Embedded Figures (number correct)	Baseline	36	2.88	1.79
	Placebo	36	3.25	1.87
	MAS	36	3.39	1.78
Raven (number correct)	Baseline	37	7.27	1.87
	Placebo	37	8.19	2.16
	MAS	37	8.11	1.84
SAT Math (number correct)	Baseline	45	12.98	5.39
	Placebo	45	13.76	6.48
	MAS	45	13.07	6.18
SAT Verbal (number correct)	Baseline	45	29.42	6.68
	Placebo	45	30.73	7.25
	MAS	45	30.29	7.51

Table 1.3.

From a total of 702 main effects and interactions, based on 78 univariate statistical analyses conducted, the following were significant but not reported in the text. The effects in this table are extraneous to our predictions, and hence not discussed further.

Task (Measure)	Statistical Test	Main Effects and Interactions	*df*	*F*	*p* uncorrected
Face Recognition (number correct)	2 (Drug:) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x Version Order	1, 36	14.15	0.001
Face Recognition (number correct)	2 (Drug) x 2 (COMT genotype: *val-val; met-met)* x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x Version Order	1, 10	6.44	0.029
Face Recognition (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among low-performers	Drug x Drug Order x Version Order	1, 19	11.51	0.003
Word Recall (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 36	8.21	0.007
Word Recognition (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Version Order	1, 36	5.06	0.030
Digit Span Backwad (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 34	15.48	<.001
Digit Span Forward (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 34	5.05	0.031
Digit Span Backwad (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Baseline x Drug Order x Version Order	1, 34	8.01	0.007
Digit Span Backwad (number correct)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x COMT	1, 10	4.97	0.050
Object-2-Back (omissions)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version	Baseline Performance	1, 37	16.37	<.001

	Order) ANOVA				
Object-2-Back (omissions)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x COMT	1, 10	4.96	0.05
Object-2-Back (omissions)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Version Order x COMT	1, 10	5.78	0.037
Go/No-go (commissions)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 34	6.87	0.013
Go/No-go (commissions)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among *val-val* participants	Drug Order x Version Order	1, 4	8.38	.044
Flanker (inhibition cost)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 35	11.89	0.001
Remote Associations (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x Version Order	1, 38	7.68	0.009
Remote Associations (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among low-performers	Drug x Drug Order x Version Order	1, 20	8.67	0.008
Embedded Figures (number correct)	2 (Drug:) x 2 (COMT genotype: *val-val; met-met)* x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order x Version Order	1,7	6.03	0.044
Embedded Figures (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among low-performers	Drug x Drug Order x Version Order	1, 12	11.35	0.006
Raven (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 29	12.54	0.001
Raven (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among low-performers	Drug x Drug Order	1, 16	5.36	0.034
Raven (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Raven Baseline x Drug Order	1, 29	6.8	0.014
Raven (number	2 (Drug) x 2 (Drug Order) x 2 (Test Version	Drug x Version Order	1, 16	8.58	0.010

correct)	Order) ANOVA among low-performers				
SAT Math (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 37	47.21	<.000
SAT Math (number correct)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	COMT x Drug Order x Version Order	1, 11	8.68	0.013
SAT Math (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Baseline x Drug Order	1, 37	4.12	0.050
SAT Math (number correct)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order	1, 11	5.29	0.042
SAT Math (number correct)	2 (Drug) x 2 (COMT genotype: val-val; met-met) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Drug x Drug Order X COMT	1, 11	9.63	0.010
SAT Math (number correct)	2 (Drug) x 2 (COMT genotype: *val-val; met-met)* x 2 (Drug Order) x 2 (Test Version Order)	Drug x Version Order x COMT	1, 11	7.9	0.017
SAT Math (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among *val-val* participants	Drug x Version Order	1, 6	6.34	.045
SAT Math (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among *val-val* participants	Drug Order x Version Order	1, 6	6.23	.042
SAT Math (number correct)	2 (Drug) x 2 (Drug Order) x 2 (Test Version Order) ANOVA among low-performers	Drug x Version Order x COMT	1, 18	6.15	0.023
SAT Verbal (number correct)	2 (Drug) x 2 (Baseline Performance) x 2 (Drug Order) x 2 (Test Version Order) ANOVA	Baseline Performance	1, 37	31.43	<.000

Table 2.1.

Eligible Measures for Examined Tasks

Cognitive Construct	Task	Eligible Measure(s)	Reference supporting choice of measure
Inhibitory control	Stop Signal Task	Depending on task design: • Stop Signal RT (Mean Go RT minus Mean Stop Delay) • Probability of inhibiting a response	• Logan et al. (1997) • Lappin & Eriksen (1966)
	Go/No-go	• False alarms or no-go accuracy	• Helmers et al. (1995), Aron et al. (2004)
	Wisconsin Card-sort	• Perseverative errors • If unavailable: accuracy	• Heaton et al. (1993)
	ID/ED	• Perseverative extra-dimensional shift errors	• Robbins et al. (1998)
	Flanker	• Difference or ratio between accuracy in the congruent and incongruent conditions • If unavailable: incongruent condition accuracy • If accuracy was at ceiling, corresponding reaction times (RTs) were coded	• Eriksen & Eriksen (1974)
	Stroop	• Difference or ratio between accuracy in the congruent and incongruent conditions • If unavailable: incongruent condition accuracy • If accuracy was at ceiling, corresponding RTs were coded	• Stroop (1935)
	Antisaccade task	• Error saccades toward the target	• Everling & Fisher (1998)
Working Memory	NBack[1]	• d', difference between hits and false alarms, or overall accuracy • If unavailable:omissions or hit rate • When the available measures from the list above were at ceiling, RTs were coded instead[2]	• Jaeggi et al. (2010), Kane et al. (2007)

	Rapid Information Processing	• Processing rate (digits presented per minute)	• Fillmore et al. (2005)
	Sternberg	• Load effect • If unavailable: accuracy • If accuracy was at ceiling, corresponding RTs were coded[2]	• Sternberg (1966)
	Digit Span	• Accuracy • If unavailable: Longest length of correctly reported item	• The Psychological Corporation (2002)
	CANTAB Spatial Working Memory	• Within and between search errors • If unavailable: Within- *or* between-search errors	• Owen et al. (1990)
	Spatial Delayed Response	• Accuracy (percent correct)	• Poste et al. (1997)
	Other WM measures	• d' or accuracy • If unavailable: omission errors • For spatial tasks: error to position and positional fit	• Jaeggi et al. (2010), Kane et al. (2007)
Immediate and delayed memory	Recall (free and cued) and recognition tests	• Sensitivity (d' or a'), proportion of hits minus proportion of false alarms, accuracy or number of trials to criterion • If unavailable: hit rate	• Henson et al. (2000)

[1] Only data from 2- and 3-Back tasks were coded, excluding data from 0-back conditions (which capitalize on sustained attention more than working memory) and 1-back conditions (which, while taxing some working memory components, such as on-line maintenance, minimally tax other facets of working memory, such as monitoring and manipulation). Thus, we only included the n-back conditions that maximized the possibility of detecting drug effect and minimized the possibility of ceiling effects.

[2] If there was no basis for inferring presence or absence of ceiling or floor effects, both accuracy and RT measures were coded.

Table 2.2.

Stimulant Enhancement of Inhibitory Control: Effect Sizes and Study Characteristics

Study	Test	Drug	Dose (mg)	*N*	Design	Dose	Caffeine Restriction	%Male	Other reason to publish?	Floor or Ceiling?	Hedges's *g*
Acheson & de Wit (2008)	Stop Signal Task	amp	20	28	Within-subjects	High	no	54.54	no	not suspected	0.23
Agay et al. (2010)	TOVA commissions	mph	15	25	Between-subjects	Low	no	46.15	no	possible	0
Allman et al. (2010)	Antisaccade Task	amp	21	24	Within-subjects	High	no	70.83	no	not suspected	0.35
Barch & Carter (2005)	Stroop	amp	17.5	22	Within-subjects	Low	no	55	no	not suspected	0.23
Costa et al. (2012)	Stop Signal Task, Go/No-go	mph	40	46	Within-subjects	High	yes	100	no	not suspected	0.07
de Bruijn et al. (2004)	Flanker	amp	15	12	Within-subjects	Low	no	58.33	no	not suspected	0.22
de Wit et al. (Unpublished)	Stop Signal Task	amp	5, 10, 20	207	Within-subjects	Both	no	52.43	no	not suspected	0.21

de Wit et al. (2000)	Stop Signal Task	amp	10, 20	20	Within-subjects	Both	no	70	no	not suspected	0.28
deWit et al. (2002)	Stop Signal Task, Go/No-go	amp	10, 20	36	Within-subjects	Both	no	50	no	not suspected	0.35
Engert et al. (2009)	WCST	mph	20	43	Between-subjects	Low	yes	100	no	not suspected	0.11
Farah et al. (Unpublished)	Flanker	amp	10	15	Within-subjects	Low	no	25	no	not suspected	0.22
Fillmore et al. (2005)	Stop Signal Task	amp	7.5, 15	22	Within-subjects	Low	yes	45.45	no	not suspected	0.1
Hamidovic et al. (2009)	Stop Signal Task	amp	5, 10, 20	93	Within-subjects	Both	no	53.76	yes	not suspected	0.2
Hester et al. (2012)	Go/No-go (modified)	mph	30	27	Within-subjects	Low	yes	100	yes	not suspected	0.18
Ilieva et al. (2013)	Go/No-go, Flanker	amp	20	43	Within-subjects	High	no	50	no	not suspected	0.05
Kelly et al. (2006)	Stop Signal Task	amp	8, 15	20	Within-subjects	Low	no	50	no	not suspected	0.09
Linssen et al. (2012)	Stop Signal Task	mph	10, 20, 40	19	Within-subjects	Both	no	100	no	not suspected	0.35
Mattay et al. (1996)	WCST	amp	17.5	8	Within-subjects	Low	yes	50	yes	not suspected	0.08
Moeller et al. (2012)	Stroop	mph	20	15	Within-subjects	Low	no	93.33	yes	not suspected	0.37

Nandam et al. (2011)	Stop Signal Task	mph	30	24	Within-subjects	Low	no	100	yes	not suspected	0.58
Pauls et al. (2012)	Stop Signal Task	mph	40	16	Within-subjects	High	yes	100	yes	not suspected	0.32
Servan-Schreiber et al. (1998)	Flanker	amp	17.5	8	Within-subjects	Low	no	50	no	not suspected	0.72
Sofuoglu et al. (2008)	Go/No-go	amp	20	10	Within-subjects	High	no	58.33	yes	possible	-0.36
Theunissen et al. (2009)	Stop Signal Task	mph	20	16	Within-subjects	Low	yes	31.25	yes	not suspected	-0.01
Overall Effect Size											0.20

Table 2.3.

Stimulant Enhancement of Working Memory: Effect Sizes and Study Characteristics

Study	Test	Drug	Dose (mg)	*N*	Design	Dose	Caffeine Restriction	%Male	Other reason to publish?	Floor or Ceiling?	Hedges's *g*
Agay et al. (2010)	Digit Span	mph	15	26	Between-subjects	Low	no	56.25	yes	not suspected	0.22
Agay et al. (Unpublished)	CANTAB Spatial WM, Digit Span	mph	20	19	Within-subjects	Low	no	not reported	yes	not suspected	0.23
Barch & Carter (2005)	Spatial Working Memory (8s delay)	amp	17.5	22	Within-subjects	Low	no	55	yes	not suspected	0.10
Dorflinger (Unpublished)	2-Back, 3-Back	mph	14, 28	20	Within-subjects	Low	no	not reported	yes	not suspected	0.15
Farah et al. (Unpublished)	2-Back, Digit Span	amp	10	16	Within-subjects	Low	no	25	no	not suspected	-0.10
Fillmore et al. (2005)	Rapid Info Processing, Spatial Delayed Resp.	amp	7.5, 15	22	Within-subjects	Low	yes	45.45	no	not suspected	0.25
Ilieva et al. (2013)	2-Back, Digit Span	amp	20	43	Within-subjects	High	no	50	no	not suspected	0.01

Kelly et al. (2006)	Rapid Info Processing, Spatial Delay Resp.	amp	7.5, 15	20	Within-subjects	Low	no	50	no	not suspected	0.38
Linssen et al. (2012)	Spatial Working Memory	mph	10, 20, 40	19	Within-subjects	Low, High	yes	100	no	not suspected	0.41
Marquand et al. (2011)	Spatial Working Memory (unrewarded)	mph	30	15	Within-subjects	Low	yes	100	yes	not suspected	-0.11
Mattay et al. (2000)	2-Back	amp	17.5	10	Within-subjects	Low	yes	80	yes	not suspected	-0.04
Mattay et al. (2003)	2-Back, 3-Back	amp	17.5	26	Within-subjects	Low	no	40.74	yes	not suspected	-0.23
Mehta et al. (2000)	CANTAB Spatial WM	mph	40	10	Within-subjects	High	no	100	yes	not suspected	0.27
Mintzer et al. (2003)	2-Back, Digit Recall	amp	20	20	Within-subjects	High	no	70	yes	possible on one measure	0.25
Mintzer et al. (2007)	2-Back, 3-Back	amp	20, 30	18	Within-subjects	High	no	61.11	yes	not suspected	0.15
Oken et al. (1995)	Digit Span	mph	14	23	Within-subjects	Low	yes	47.83	yes	not suspected	-0.14
Ramasubbu et al. (2012)	2-Back	mph	20	13	Within-subjects	Low	yes	38.46	yes	not suspected	0.62
Schmedtje et al.	Pattern Memory,	amp	5	8	Within-	Low	no	not	yes	not	0.30

(1988)	Digit Span				subjects			reported		suspected	
Silber et al. (2006)	Digit Span	amp	5	20	Within-subjects	Low	yes	50	yes	not suspected	0.18
Studer et al. (2010)	Spatial Working Memory	mph	20	11	Within-subjects	Low	no	45.45	yes	possible on one measure	0.10
Overall Effect Size											0.13

Table 2.4.

Stimulant Enhancement of Short-Term Episodic Memory: Effect Sizes and Study Characteristics

Study	Test	Drug	Dose (mg)	*N*	Design	Dose	Caffeine Restriction	Retention Interval	%Male	Other reason to publish?	Floor or Ceiling?	*g*
Farah et al. (Unpublished)	Word Recall, Word Recognition, Face Recognition	amp	10	16	Within-subjects	Low	no	30 min	25	no	not suspected	0.07
Fleming et al. (1995)	Paired Associates, Rey Verbal Learning Test	amp	20	17	Within-subjects	High	no	immediate	52.94	no	possible on one measure	0.16
Linssen et al. (2012)	Word Recall, Word Recognition	mph	10, 20, 40	19	Within-subjects	Low, High	no	30 min, immediate	100	no	possible on one measure	0.20
Soetens et al. (1995), Study 1	Word Recall	amp	10	18	Within-subjects	Low	no	immediate	100	no	not suspected	0.23
Soetens et al. (1995), Study 2	Word Recall	amp	10	14	Within-subjects	Low	no	immediate	100	no	not suspected	0.39
Soetens et al. (1995), Study 4	Word Recognition	amp	10	12	Within-subjects	Low	no	immediate	100	no	not suspected	0.29
Soetens et al. (1995), Study 5	Word Recognition	amp	10	12	Within-subjects	Low	no	immediate	100	no	not suspected	0.31

Unrug et al. (1997)	Word Recall	mph	20	12	Within-subjects	Low	yes	20 min	50	yes	possible on one measure	0.32
Willet (1962)	Learning of Non-word Lists	amp	10	37	Between-subjects	Low	no	immediate	0	no	not suspected	0.22
Zeeuws et al. (2010b), Exp. 1	Word Recognition	amp	10	24	Within-subjects	Low	no	immediate	100	no	not suspected	-0.17
Zeeuws et al. (2010b), Exp. 2	Word Recognition	amp	10	16	Within-subjects	Low	no	immediate	100	no	not suspected	-0.13
Zeeuws & Soetens (2007)	Word Recognition	amp	10	36	Within-subjects	Low	no	30 min, immediate	100	no	not suspected	0.45
Overall Effect Size												0.20

Table 2.5.

Stimulant Enhancement of Long-Term Episodic Memory: Effect Sizes and Study Characteristics

Study	Test	Drug	Dose (mg)	*N*	Design	Dose	Caffeine Restriction	Retention Interval	%Male	Other reason to publish?	Floor or Ceiling?	*g*
Brignell et al. (2007)	Recognition Memory for Narratives	mph	40	36	Between-subjects	High	no	1 hour, 1 day	not reported	yes	not suspected	0.52
Ilieva et al. (2013)	Word Recall, Word Recognition, Face Recognition	amp	20	18	Within-subjects	High	no	2 hours	50	no	not suspected	0.01
Mintzer et al. (2003)	Word Recall and Recognition	amp	20	16	Within-subjects	High	no	2 hours	70	yes	not suspected	0.24
Mintzer et al. (2007)	Word Recall and Recognition	amp	20, 30	20	Within-subjects	High	no	2 hours	61.11	yes	not suspected	0.33
Soetens et al. (1995), Exp.1	Word Recall	amp	10	44	Within-subjects	Low	no	1 day	100	no	not suspected	0.71

Soetens et al. (1995), Exp.2	Word Recall	amp	10	18	Within-subjects	Low	no	1 hour, 1 day	100	no	not suspected	0.58
Soetens et al. (1995), Exp.4	Word Recall	amp	10	14	Within-subjects	Low	no	1 day, 2 days, 3 days	100	no	not suspected	0.58
Soetens et al. (1995), Exp.5	Word Recognition	amp	10	12	Within-subjects	Low	no	1 day, 1 week	100	no	not suspected	0.74
Zeeuws et al. (2010b), Exp.1	Word Recognition	amp	10	12	Within-subjects	Low	no	1 hour, 1 day, 1 week	100	no	not suspected	0.69
Zeeuws et al. (2010), Exp.2	Word Recognition	amp	10	24	Within-subjects	Low	no	1 hour, 1 day, 1 week	100	no	not suspected	0.18
Zeeuws & Soetens (2007)	Word Recognition	amp	10	16	Within-subjects	Low	no	1 hour, 1 day	100	no	not suspected	0.80
Overall Effect Size												0.45

FIGURE CAPTIONS

Figure 1.1. Embedded Figures Task: An example stimulus.

Figure 1.2. Experimental procedure. Each box corresponds to an individual testing session, with the time intervals between sessions indicated. Baseline testing (Sessions 2-3) always preceded drug/placebo testing (Sessions 4-7) to minimize the influence of practice effects on data from the placebo and Adderall conditions. Each individual participant's four on-pill sessions were scheduled at the same time of the day.

Figure 1.3. Mean performance of participants whose overall baseline performance was below and above the median on Word Recall, Embedded Figures and Raven's Progressive Matrices. Conventional error bars are not shown because the comparisons are within subjects.

Figure 1.4. Mean performance of participants homozygous for the *val* and *met* allele of the COMT gene on SAT Math and SAT Verbal. Conventional error bars are not shown because the comparisons are within subjects.

Figure 2.1. Process of determining study eligibility

Figure 2.2. Funnel plot of research on stimulant effects on inhibitory control. Data points imputed by trim and fill appear in black.

Figure 2.3. Funnel plot of research on stimulant effects on working memory. Data points imputed by trim and fill appear in black.

Figure 2.4. Funnel plot of research on stimulant effects on short-term episodic memory. Data points imputed by trim and fill appear in black.

Figure 2.5. Funnel plot of research on stimulant effects on long-term episodic memory. Data points imputed by trim and fill appear in black.

Figure 3.1. Discrepancy between subjectively reported and objectively measured attention in users of unprescribed stimulants and controls, based on data on: a) overall attention tests; b) inattention subtests; c) impulsivity subtests.

FIGURES

Figure 1.1

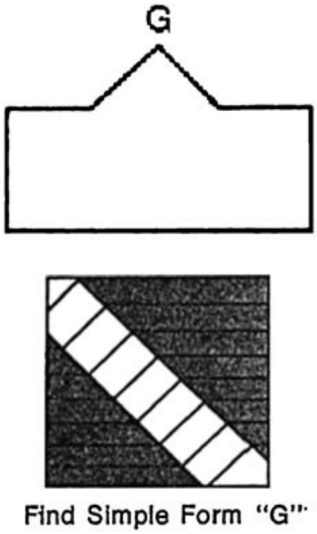


Figure 1.2.

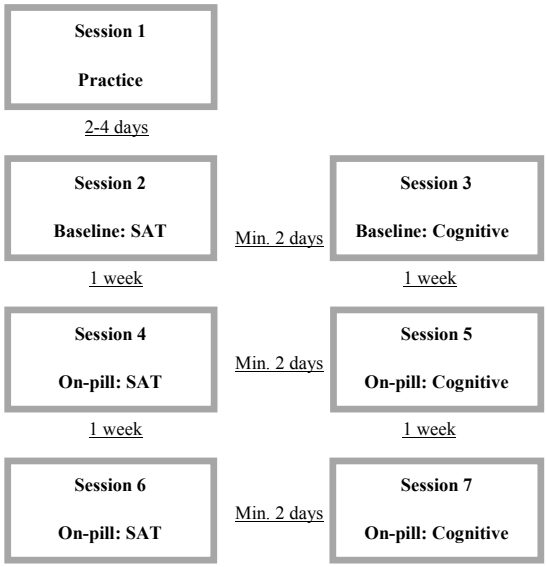
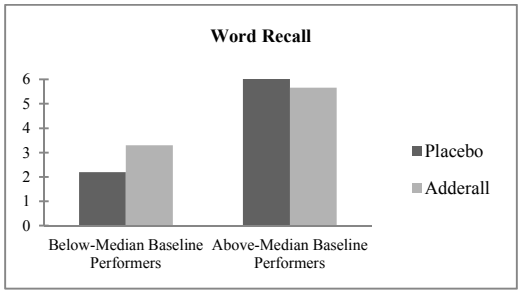
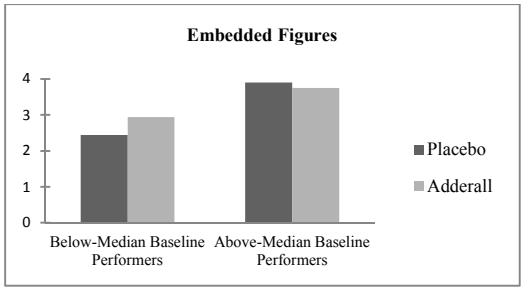

Session 1
Practice
2-4 days
Session 2
Baseline: SAT
Min. 2 days
Session 3
Baseline: Cognitive
1 week
1 week
Session 4
On-pill: SAT
Min. 2 days
Session 5
On-pill: Cognitive
1 week
1 week
Session 6
On-pill: SAT
Min. 2 days
Session 7
On-pill: Cognitive

Figure 1.3.


Word Recall
6
5
4
3
2
1
0
Below-Median Baseline Performers
Above-Median Baseline Performers
Placebo
Adderall


Embedded Figures
4
3
2
1
0
Below-Median Baseline Performers
Above-Median Baseline Performers
Placebo
Adderall

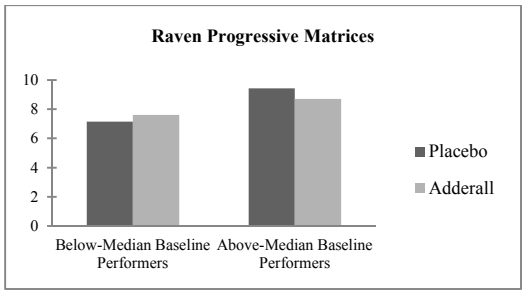
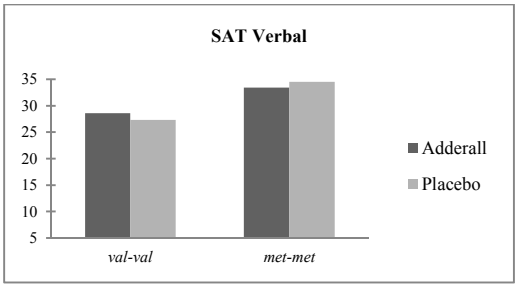

Raven Progressive Matrices
10
8
6
4
2
0
Below-Median Baseline Performers
Above-Median Baseline Performers
Placebo
Adderall

Figure 1.4.


SAT Verbal
35
30
25
20
15
10
5
val-val
met-met
Adderall
Placebo

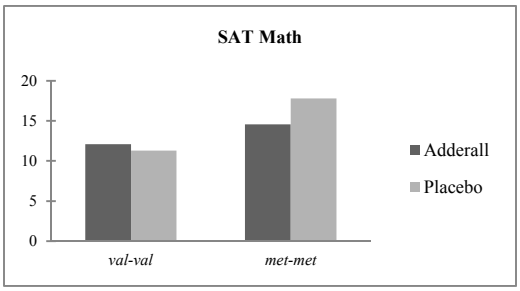

SAT Math
20
15
10
5
0
val-val
met-met
Adderall
Placebo

Figure 2.1.

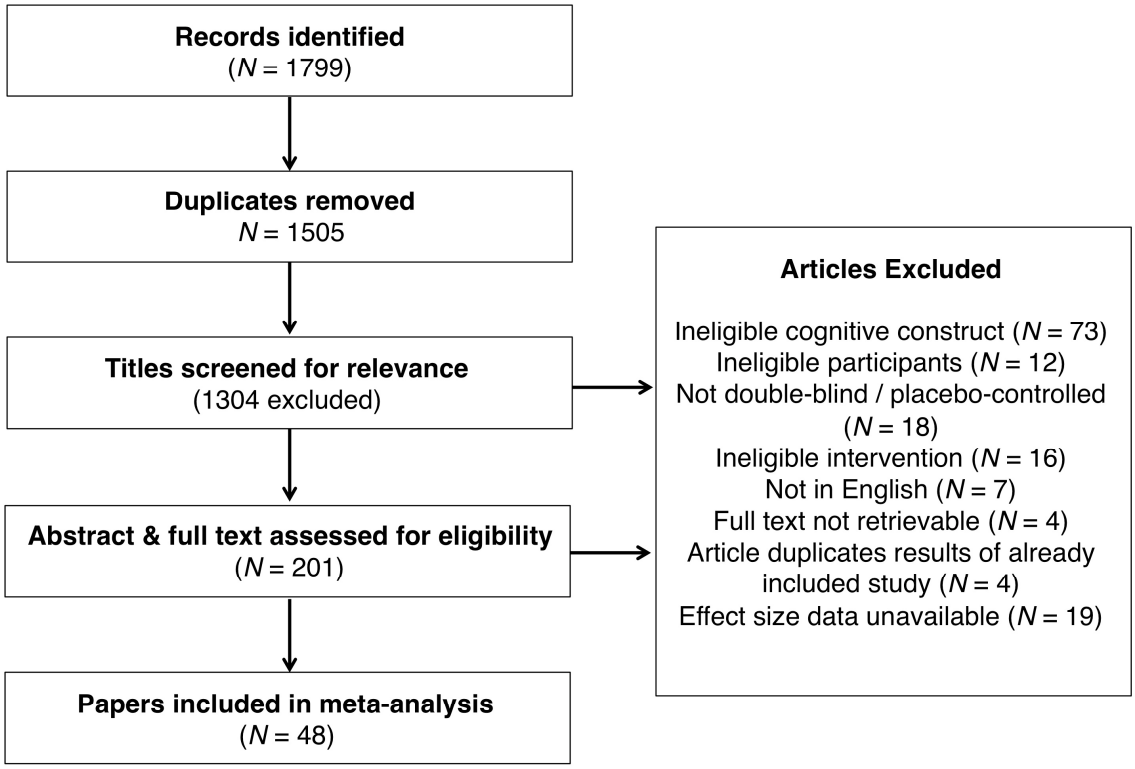

Records identified
(N = 1799)
Duplicates removed
N = 1505
Titles screened for relevance
(1304 excluded)
Abstract & full text assessed for eligibility
(N = 201)
Papers included in meta-analysis
(N = 48)
Articles Excluded
Ineligible cognitive construct (N = 73)
Ineligible participants (N = 12)
Not double-blind / placebo-controlled (N = 18)
Ineligible intervention (N = 16)
Not in English (N = 7)
Full text not retrievable (N = 4)
Article duplicates results of already included study (N = 4)
Effect size data unavailable (N = 19)

Figure 2.2.

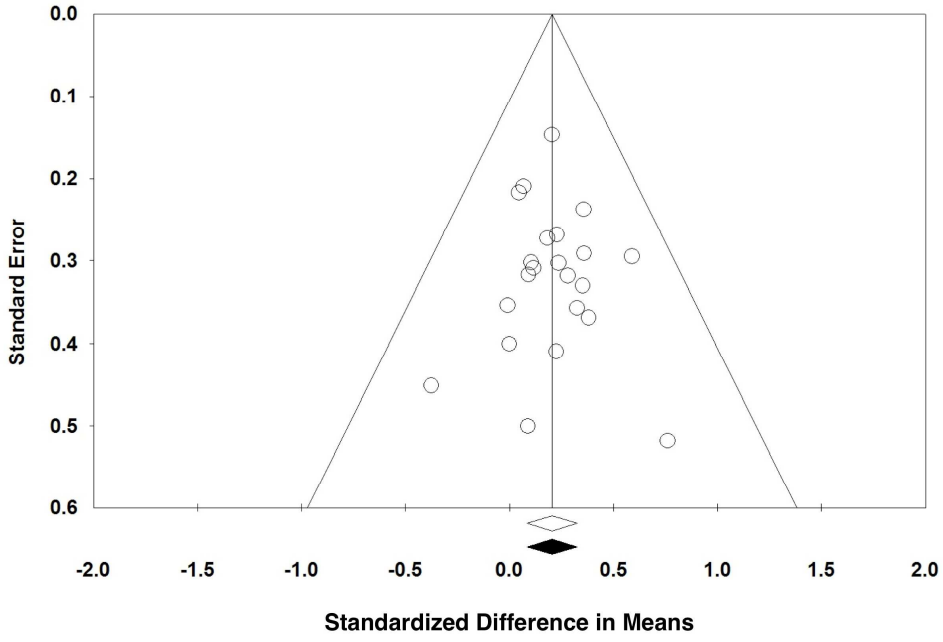

0.0
0.1
0.2
0.3
0.4
0.5
0.6
Standard Error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Standardized Difference in Means

Figure 2.3.

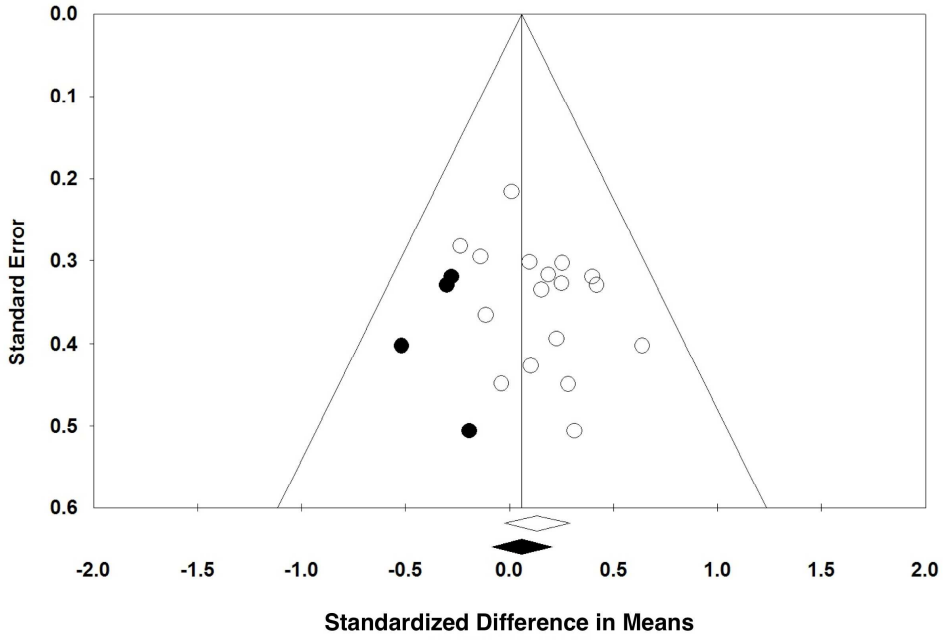

0.0
0.1
0.2
0.3
0.4
0.5
0.6
Standard Error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Standardized Difference in Means

Figure 2.4.

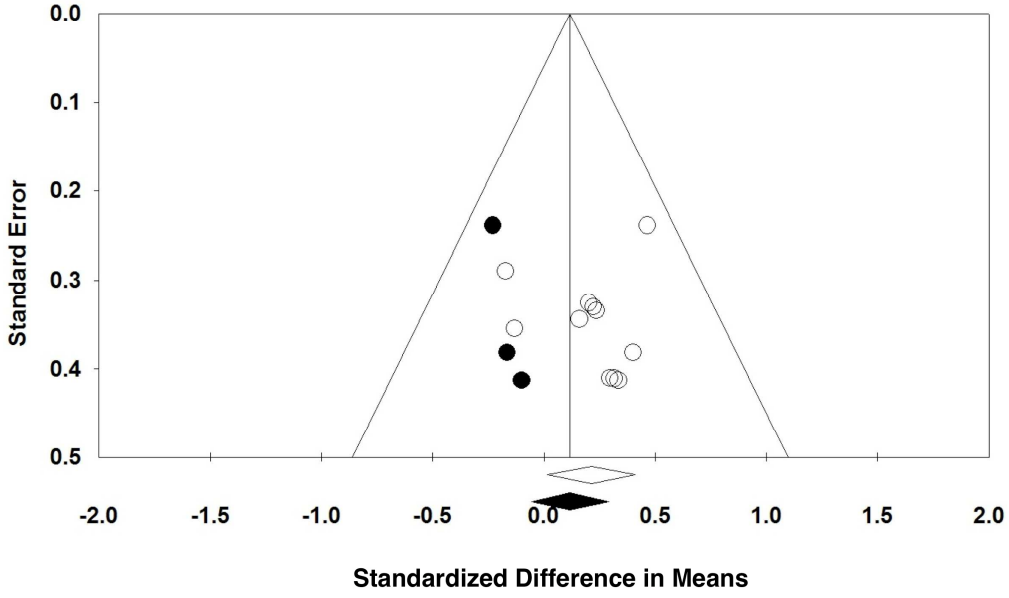

0.0
0.1
0.2
0.3
0.4
0.5
Standard Error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Standardized Difference in Means

Figure 2.5.

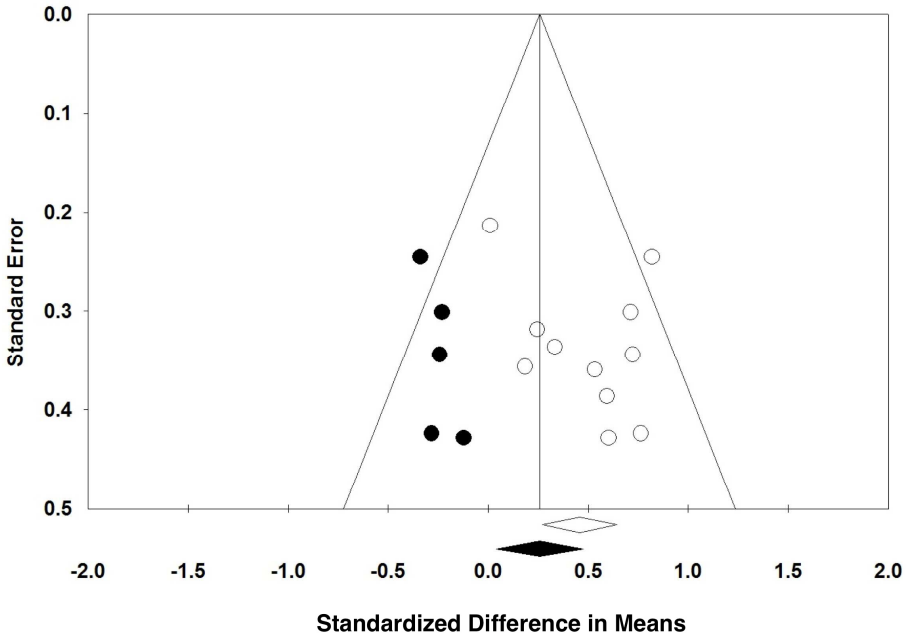

0.0
0.1
0.2
0.3
0.4
0.5
Standard Error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Standardized Difference in Means

Figure 3.1. (a)

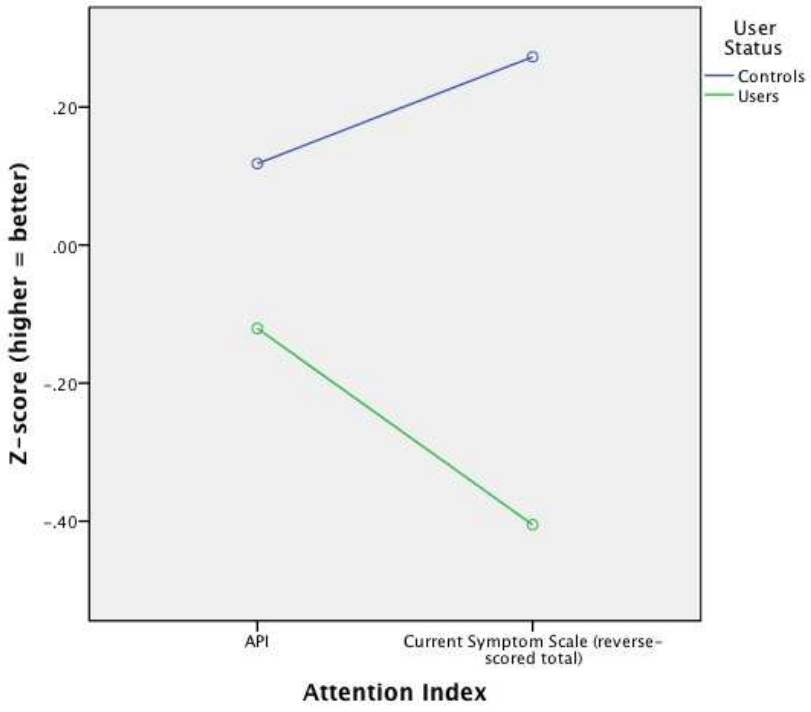

User Status
Controls
Users
.20
.00
-.20
-.40
Z-score (higher = better)
API
Current Symptom Scale (reverse-scored total)
Attention Index

Figure 3.1. (b)

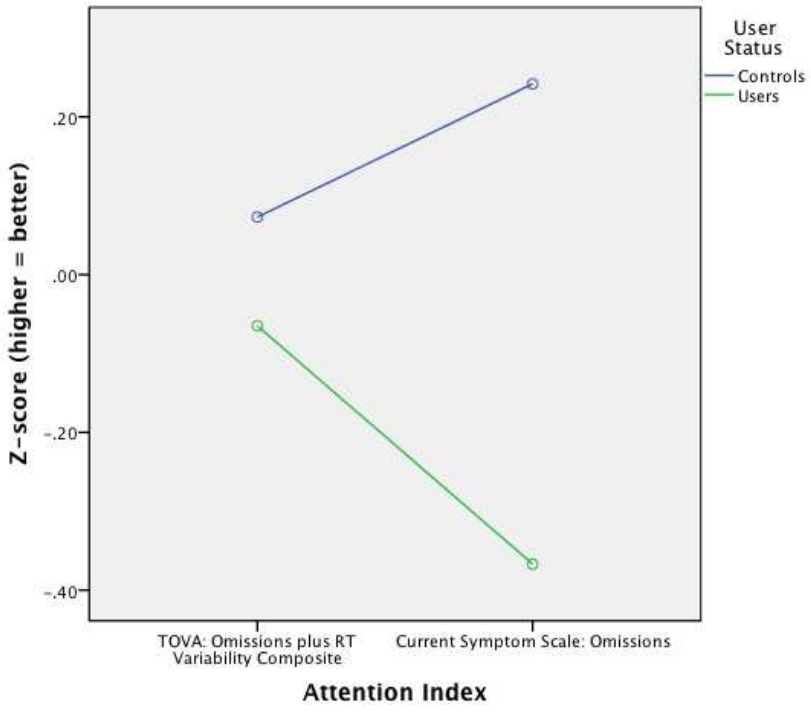
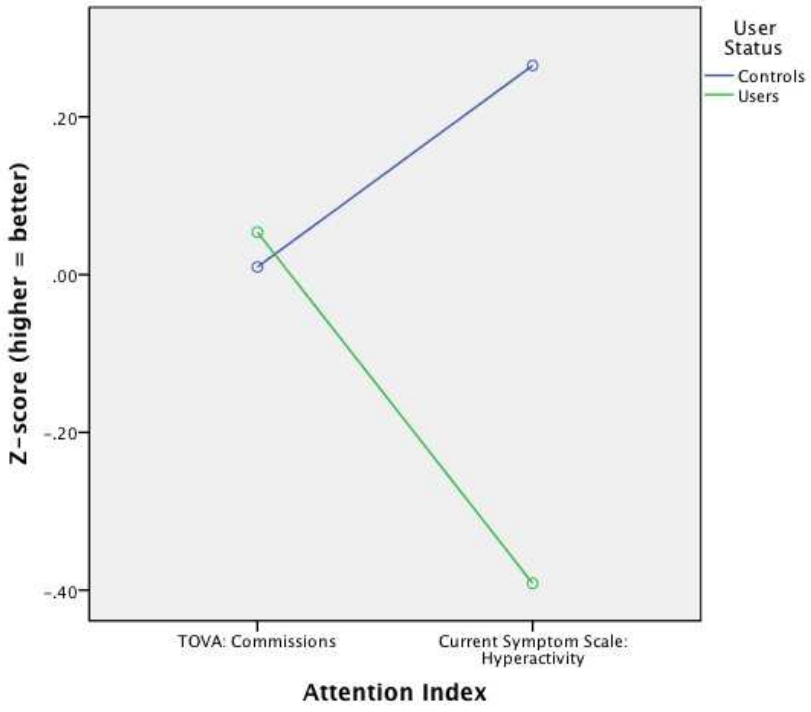

User Status
Controls
Users
.20
.00
-.20
-.40
Z-score (higher = better)
TOVA: Omissions plus RT Variability Composite
Current Symptom Scale: Omissions
Attention Index

Figure 3.1. (c)


User Status
Controls
Users
.20
.00
-.20
-.40
Z-score (higher = better)
TOVA: Commissions
Current Symptom Scale: Hyperactivity
Attention Index

BIBLIOGRAPHY

Acheson, A., & de Wit, H. (2008). Bupropion improves attention but does not affect impulsive behavior in healthy young adults. *Experimental and Clinical Psychopharmacology*, *16*(2), 113–123.

Advokat, C. (2010). What are the cognitive effects of stimulant medications? Emphasis on adults with attention-deficit/hyperactivity disorder (ADHD). *Neuroscience & Biobehavioral Reviews*, *34*(8), 1256–1266.

Agay, N. (2012). [Cognitive Effects of Stimulants.] Unpublished raw data.

Agay, N., Yechiam, E., Carmel, Z., & Levkovitz, Y. (2010). Non-specific effects of methylphenidate (Ritalin) on cognitive ability and decision-making of ADHD and healthy adults. *Psychopharmacology*, *210*(4), 511–519.

Allman, A.-A., Benkelfat, C., Durand, F., Sibon, I., Dagher, A., Leyton, M., … O'Driscoll, G. A. (2010). Effect of d-amphetamine on inhibition and motor planning as a function of baseline performance. *Psychopharmacology*, *211*(4), 423–433.

Angrist, B., Corwin, J., Bartlik, B., Cooper, T. (1987). Early pharmacokinetics and clinical effects of oral D-amphetamine in normal subjects. *Biol psychiatry*, *22*(11), 1357-68.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*(2), 115–116.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Arria, A. M., Garnier-dykstra, L. M., Caldeira, K. M., Vincent, K. B., Grady, K. E. O., & Wish, E. D. (2011). Non-medical use of Prescription Stimulants Among College Students : Possible Association With ADHD Symptoms. doi:10.1177/1087054710367621

Bados, A., Gómez-Benito, J., & Balaguer, G. (2010). The state-trait anxiety inventory, trait version: does it really measure anxiety? *Journal of personality assessment*, 92(6), 560–7. doi:10.1080/00223891.2010.513295

Ballard, M. E., Gallo, D. a, & de Wit, H. (2014). Amphetamine increases errors during episodic memory retrieval. *Journal of clinical psychopharmacology*, 34(1), 85–92. doi:10.1097/JCP.0000000000000039

Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L., & Saggino, A. (2013). The State-Trait Anxiety Inventory: Shadows and Lights on its Construct Validity. *Journal of Psychopathology and Behavioral Assessment*, 35(4), 475–486. doi:10.1007/s10862-013-9354-5

Baranski, J. V., Pigeau, R. A. (1997). Self-monitoring cognitive performance during sleep deprivation: effects of modafinil, d-amphetamine and placebo. *J Sleep Res*, *6*, 84-91.

Baranski, Joseph V, Pigeau, R., Dinich, P., Jacobs, I. (2004). Effects of modafinil on cognitive and meta-cognitive performance. *Human Psychopharm*, *19*, 323-32.

Barch, D. M., & Carter, C. S. (2005). Amphetamine improves cognitive function in medicated individuals with schizophrenia and in healthy volunteers. *Schizophrenia Research*, *77*(1), 43–58.

Bavarian, N., Flay, B. R., Ketcham, P. L., and Smit, E. (2013). Illicit use of prescription stimulants in a college student sample: a theory-guided analysis. *Drug Alcohol Depend*. 132, 665–673. doi: 10.1016/j.drugalcdep.2013. 04.024

Biggs, J. (1987). *Study Process Questionnaire Manual: Student Approaches to Learning and Studying*. Hawthorn, Australia: Australian Council for Education Research.

Bjork, R. a, Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: beliefs, techniques, and illusions. *Annual review of psychology*, 64, 417–44. doi:10.1146/annurev-psych-113011-143823

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb cortex*, *11*, 825-36.

Brignell, C. M., Rosenthal, J., & Curran, H. V. (2007). Pharmacological manipulations of arousal and memory for emotional material: effects of a single dose of methylphenidate or lorazepam. *Journal of Psychopharmacology*, *21*(7), 673–683.

Bruijn, E. R. A. de, Hulstijn, W., Verkes, R. J., Ruigt, G. S. F., & Sabbe, B. G. C. (2004). Drug-induced stimulation and suppression of action monitoring in healthy volunteers. *Psychopharmacology*, *177*(1-2), 151–160.

Burns, J. T., House, R. F., Fensch, F. C., & Miller, J. G. (1967). Effects of Magnesium Pemoline and Dextroamphetamine on Human Learning. *Science*, *155*(3764), 849–851.

Butcher, S. P., Fairbrother, I. S., Kelly, J. S., and Arbuthnott, G. W. (1988).

Amphetamine- induced dopamine release in the rat striatum. *J. Neurochem.* 50, 346–355. doi: 10.1111/j.1471-4159. 1988.tb02919.x

Castaldi, S., Gelatti, U., Orizio, G., Hartung, U., Moreno-londono, A. M., Nobile, M., et al. (2012). Use of cognitive enhancement medication among Northern Italian university students. *J Addict. Med*. 6, 112–117. doi: 10.1097/ADM.0b013e3182479584

Chamberlain, S. R., Robbins, T. W., Winder-Rhodes, S., Muller, U., Sahakian, B. J., Blackwell, A. D., Barnett, J. H. (2010). Translational approaches to frontostriatal dysfunction in attention-deficit/hyperactivity disorder using a computerized neuropsychological battery. *Biol Psychiatry*, *69*, 1192–1203.

Chatterjee, A. (2009). Is it acceptable for people to take methylphenidate to enhance performance ? No Is it acceptable for people to take methylphenidate to enhance performance ?, (June), 6–8. doi:10.1136/bmj.b1956

Christopoulos, J. P., Rohwer, W. D., & Thomas, J. W. (1987). Grade level differences in students' study activities as a function of course characteristics. *Contemporary Educational Psychology*, 12(4), 303–323. doi:10.1016/S0361-476X(87)80003-6

Cooper, H. (2010). *Research Synthesis and Meta-analysis.* Thousand Oaks, CA: SAGE Publications, Inc.

Costa, A., Riedel, M., Pogarell, O., Menzel-Zelnitschek, F., Schwarz, M., Reiser, M., … Ettinger, U. (2013). Methylphenidate Effects on Neural Activity During Response Inhibition in Healthy Humans. *Cerebral Cortex*, *23*(5), 1179–1189.

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, 21(4), 337–346. doi:10.1016/j.lindif.2011.03.002

Davis, D. R., (1947). Psychomotor effects of analeptics and their relation to fatigue phenomena in aircrew. *Brit. Med. Bull., 5*, 43-45.

De Wit, H. (2012). [Stimulants' Effects on Stop Signal Task Performance.] Unpublished raw data.

De Wit, H., Crean, J., & Richards, J. B. (2000). Effects of d-Amphetamine and ethanol on a measure of behavioral inhibition in humans. *Behavioral Neuroscience*, *114*(4), 830–837.

De Wit, H., Enggasser, J. L., & Richards, J. B. (2002). Acute Administration of d-Amphetamine Decreases Impulsivity in Healthy Volunteers. *Neuropsychopharmacology*, *27*(5), 813–825.

DeSantis, A. D., & Hane, A. C. (2010). "Adderall is definitely not a drug": justifications for the illegal use of ADHD stimulants. *Substance use & misuse*, 45(1-2), 31–46. doi:10.3109/10826080902858334

DeSantis, A. D., Webb, E. M., and Noar, S. M. (2008). Illicit use of prescription ADHD medications on a college campus: a multi- methodological approach. *J. Am. College Health* 57, 315–324. doi: 10.3200/JACH.57.3.315-324

Desantis, A., Noar, S. M., Webb, E. M. (2009). Nonmedical ADHD stimulant use in fraternities. *J of studies on alcohol and drugs*, *70*(6), 952-4.

Dorflinger, J. (2005). *Methylphenidate and working memory: An investigation using fMRI* (unpublished dissertation). Rosalind Franklin University of Medicine and Science, Chicago, IL.

Drevets, W. C., Gautier, C., Price, J. C., Kupfer, D. J., Kinahan, P. E., Grace, A. A., et al. (2001). Amphetamine-induced dopamine release in human ventral stria- tum correlates with euphoria. *Biol. Psychiatry* 49, 81–96. doi: 10.1016/S0006-3223(00)01038-6

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177.

Dunlosky, J., Rawson, K. a., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. doi:10.1177/1529100612453266

Dussault, C. L., & Weyandt, L. L. (2013). An examination of prescription stimulant misuse and psychological variables among sorority and fraternity college populations. *Journal of attention disorders*, 17(2), 87–97. doi:10.1177/1087054711428740

Elliott, R., Sahakian, B. J., Matthews, K., Bannerjea, A., Rimmer, J., Robbins, T. W., (1997). Effects of methylphenidate on spatial working memory and planning in healthy young adults. *Psychopharmacology, 131*, 196-206.

Emanuel, R. M., Frellsen, S. L., Kashima, K. J., Sanguino, S. M., Sierles, F. S., & Lazarus, C. J. (n.d.). Cognitive Enhancement Drug Use Among Future Physicians :

Findings from a Multi-Institutional Census of Medical Students, 9, 1028–1034. doi:10.1007/s11606-012-2249-4

Engert, V., Joober, R., Meaney, M. J., Hellhammer, D. H., & Pruessner, J. C. (2009). Behavioral response to methylphenidate challenge: Influence of early life parental care. *Developmental Psychobiology*, *51*(5), 408–416.

Epstein, J. N., Goldberg, N. a., Conners, C. K., & March, J. S. (1997). The effects of anxiety on continuous performance test functioning in an ADHD clinic sample. *Journal of Attention Disorders*, 2(1), 45–52. doi:10.1177/108705479700200106

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Everling, S., & Fischer, B. (1998). The antisaccade: a review of basic research and clinical studies. *Neuropsychologia*, *36*(9), 885–899.

Farah, M. J. (2012). [Stimulants' Cognitive Enhancement Effects in Healthy People.] Unpublished raw data.

Farah, M. J., Illes, J., Cook-Deegan, R., Gardner, H., Kandel, E., King, P., Parens, E., et al. (2004). Neurocognitive enhancement: what can we do and what should we do? *Nature reviews. Neuroscience*, 5(5), 421–5. doi:10.1038/nrn1390

Farah, M.J., Haimm, C., Sankoorikal, G., Smith, M.E., Chatterjee, A. (2008). When we enhance cognition with Adderall do we sacrifice creativity? A preliminary study. *Psychopharmacology,* 202, 541-547.

Fillmore, M. T., Kelly, T. H., & Martin, C. A. (2005). Effects of d-amphetamine in human models of information processing and inhibitory control. *Drug and Alcohol Dependence*, *77*(2), 151–159.

Fleming, K., Bigelow, L. B., Weinberger, D. R., & Goldberg, T. E. (1995). Neuropsychological effects of amphetamine may correlate with personality characteristics. *Psychopharmacology Bulletin*, *31*(2), 357–362.

Forbes, G. B. (1998). Clinical utility of the test of variables of attention (TOVA) in the diagnosis of attention- deficit/hyperactivity disorder. Journal of clinical psychology, 54(4), 461-476.

Franke, A. G. (2011). Non-medical use of prescription stimulants and illicit use of stimulants for cognitive enhancement in pupils and students in Germany. *Pharmacopsychiatry* 44, 60–66. doi: 10.1055/s-0030-1268417

Franke, A. G., Bagusat, C., Dietz, P., Hoffmann, I., Simon, P., Ulrich, R., & Lieb, K. (2013). Use of illicit and prescription drugs for cognitive or mood enhancement among surgeons. *BMC medicine*, 11, 102. doi:10.1186/1741-7015-11-102

Gredler, M.E. & Garavalia, L.S. (2000). Students' perceptions of their self-regulatory and other-directed study strategies: A factor analysis. *Psychological Reports,* 86, 102–108.

Greely, H. T. (2013). Some First Steps Toward Responsible Use of Cognitive-Enhancing Drugs by the Healthy. *The American Journal of Bioethics*, *13*(7), 39–41.

Greenberg, L. M., & Waldman, I. D. (1993). Developmental normative data on the Test of Variables of Attention (T. O. V. A.). *Journal of Child Psychology and*

Psychiatry, 34, 1019 –1030. doi:10.1111/j.1469- 7610.1993.tb01105.x

Greenberg, L. M., Kindschi, C. L., Dupuy, T. R., & Hughes, S. J. (2007). *TOVA clinical manual*. Los Alamitos, CA: TOVA Company.

Hall, W. D., & Lucke, J. C. (2010). The enhancement use of neuropharmaceuticals: more scepticism and caution needed. *Addiction*, *105*(12), 2041–2043.

Hamidovic, A., Dlugos, A., Skol, A., Palmer, A. A., & de Wit, H. (2009). Evaluation of genetic variability in the dopamine receptor D2 in relation to behavioral inhibition and impulsivity/sensation seeking: An exploratory study with d-amphetamine in healthy participants. *Experimental and Clinical Psychopharmacology*, *17*(6), 374–383.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). Wisconsin Card Sorting Test (WCST) Manual (Revised and Expanded). Odessa: Psychological Assessment Resources Inc.

Helmers, K. F., Young, S. N., & Pihl, R. O. (1995). Assessment of measures of impulsivity in healthy male volunteers. *Personality and Individual Differences*, *19*(6), 927–935.

Henson, R. N. A., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in Recognition Memory for Words: Dissociating Right Prefrontal Roles in Episodic Retrieval. *Journal of Cognitive Neuroscience*, *12*(6), 913–923.

Hester, R., Nandam, L. S., O'Connell, R. G., Wagner, J., Strudwick, M., Nathan, P. J., … Bellgrove, M. A. (2012). Neurochemical Enhancement of Conscious Error Awareness. *The Journal of Neuroscience*, *32*(8), 2619–2627.

Hildt, E., Lieb, K., & Franke, A. G. (2014). Life context of pharmacological academic performance enhancement among university students--a qualitative approach. *BMC medical ethics*, 15, 23. doi:10.1186/1472-6939-15-23

Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology*, *80*(2), 327–334.

Hunt, M. G., Momjian, A. J., & Wong, K. K. (2011). Effects of diurnal variation and caffeine consumption on Test of Variables of Attention (TOVA) performance in healthy young adults. *Psychological assessment*, 23(1), 226–33. doi:10.1037/a0021401

Hurst, P.M., Weidner, M., Radlow, R., (1967). The effects of amphetamines upon judgments and decisions. *Psychopharmacology.* 11, 397-404.

Ilieva, I. & Farah, M.J. (2013). Enhancement stimulants: Perceived motivational and cognitive advantages. *Frontiers in Neuroscience – Neuropharmacology,* 7 *(1).*

Ilieva, I., Boland, J., & Farah, M. J. (2013). Objective and subjective cognitive enhancing effects of mixed amphetamine salts in healthy people. *Neuropharmacology*, *64*, 496–505.

Ioannidis J. P. (2005). Why most published research findings are false. *PLoS Med* **2** (8).

Ioannidis J. P. (2011). Excess significance bias in the literature on brain volume abnormalities. *Arch Gen Psychiatry*. 68(8), 773-80.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412.

Johnson, L. G., Ph, D., & Benedict, K. B. (2006). Neurocognition in Depression : Patients on and, 217–225.

Kane, M. J., A, R., Miura, T. K., & H, J. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615–622.

Kelly, T. H., Robbins, G., Martin, C. A., Fillmore, M. T., Lane, S. D., Harrington, N. G., & Rush, C. R. (2006). Individual differences in drug abuse vulnerability: d-Amphetamine and sensation-seeking status. *Psychopharmacology*, *189*(1), 17–25.

Lappin, J. S., & Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology*, *72*(6), 805–811.

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). Evidence based medicine: The case of the misleading funnel plot. *BMJ: British Medical Journal*, *333*(7568), 597–600.

Lehrer, J. (2010). The Truth Wears Off: Is there something wrong with the scientific method? The New Yorker, Retrieved from: http://www.newyorker.com/reporting/2010/12/13/101213fa_fact_lehrer#ixzz1zsJhcmej

Linssen, A. M. W., Vuurman, E. F. P. M., Sambeth, A., & Riedel, W. J. (2012). Methylphenidate produces selective enhancement of episodic memory consolidation in healthy volunteers. *Psychopharmacology*, *221*(4), 611–619.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis.* Thousand Oaks, CA: Sage.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and Inhibitory Control. *Psychological Science*, *8*(1), 60–64.

Low, K. G., & Gendaszek, A. E. (2002). Illicit use of psychostimulants among college students: A preliminary study. *Psychology, Health & Medicine, 7,* 283–287. doi:10.1080/13548500220139386

Madrigal, A. (2008). 20 percent of scientists admit using brainenhancing drugs—do you? *Wired.* Retrieved from http://www.wired.com/wiredscience/2008/04/20-of-scientist/

Maher, B., (2008). Poll results: look who's doping. Nature 452, 674–675.

Marquand, A. F., De Simoni, S., O'Daly, O. G., Williams, S. C., Mourão-Miranda, J., & Mehta, M. A. (2011). Pattern Classification of Working Memory Networks Reveals Differential Effects of Methylphenidate, Atomoxetine, and Placebo in Healthy Volunteers. *Neuropsychopharmacology*, *36*(6), 1237–1247.

Mattay, V. S., Berman, K. F., Ostrem, J. L., Esposito, G., Horn, J. D. V., Bigelow, L. B., & Weinberger, D. R. (1996). Dextroamphetamine Enhances "Neural Network-Specific" Physiological Signals: A Positron-Emission Tomography rCBF Study. *The Journal of Neuroscience*, *16*(15), 4816–4822.

Mattay, V. S., Callicott, J. H., Bertolino, A., Heaton, I., Frank, J. A., Coppola, R., (2000). Effects of dextroamphetamine on cognitive performance and cortical activation. *NeuroImage, 12*, 268-275.

Mattay, V. S., Goldberg, T. E., Fera, F., Hariri, A. R., Tessitore, A., Egan, M. F., Kolachana, B.,, Callicott, J.H., Weinberger, D.R., (2003). Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. *PNAS, 100*, 6186-6191.

McCabe, S. E., Knight, J. R., Teter, C. J., Wechsler, H. (2005). Non-medical use of prescription stimulants among US college students: Prevalence and correlates from a national survey. *Addiction, 100*, 96-106.

McCabe, S. E., Teter, C. J., & Boyd, C. J. (2006). Medical use, illicit use, and diversion of prescription stimulant medication. *Journal of Psychoactive Drugs, 38,* 43–56.

McCabe, S. E. (2008). Misperceptions of non-medical prescription drug use: a web survey of college students. *Addictive behaviors*, 33(5), 713–24. doi:10.1016/j.addbeh.2007.12.008

McCabe, S. E., & West, B. T. (2013). Medical and nonmedical use of prescription stimulants: results from a national multicohort study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 52(12), 1272–80. doi:10.1016/j.jaac.2013.09.005

Mednick, S., (1962). The associative basis of the creative process. *Psychol Rev,* 69 (3).

Mehta, M. A., Owen, a M., Sahakian, B. J., Mavaddat, N., Pickard, J. D., Robbins, T. W., (2000). Methylphenidate enhances working memory by modulating discrete frontal and parietal lobe regions in the human brain. *J Neurosci, 20*(6).

Meyer, J.S., Quenzer, L.F. (2005). Psychopharmacology: Drugs, the Brain, and Behavior. Sunderland,MA: Sinauer Associates, Inc.

Mintzer, M. Z., & Griffiths, R. R. (2003). Triazolam-amphetamine interaction: dissociation of effects on memory versus arousal. *Journal of Psychopharmacology*, *17*(1), 17–29.

Mintzer, M. Z., & Griffiths, R. R. (2007). A triazolam/amphetamine dose–effect interaction study: dissociation of effects on memory versus arousal. *Psychopharmacology*, *192*(3), 425–440.

Moeller, S. J., Honorio, J., Tomasi, D., Parvaz, M. A., Woicik, P. A., Volkow, N. D., & Goldstein, R. Z. (2012). Methylphenidate Enhances Executive Function and Optimizes Prefrontal Function in Both Health and Cocaine Addiction. *Cerebral Cortex*, 34 (5).

Murphy, K. R., & Adler, L. A. (2004). Assessing attention-deficit/ hyperactivity disorder in adults: Focus on rating scales. *Journal of Clinical Psychiatry, 65*(Suppl3)*,* 8–11.

Nandam, L. S., Hester, R., Wagner, J., Cummins, T. D. R., Garner, K., Dean, A. J., … Bellgrove, M. A. (2011). Methylphenidate But Not Atomoxetine or Citalopram Modulates Inhibitory Control and Response Time Variability. *Biological Psychiatry*, *69*(9), 902–904.

Noppe, L.D., Gallagher, J.M. (1977). A cognitive style approach to crative thought. *J Pers Assessment,* 41 (1).

Novak, S. P., Kroutil, L. A., Williams, R. L., & Van Brunt, D. L. (2007). The

nonmedical use of prescription ADHD medications: Results from a national Internet panel. *Substance Abuse Treatment, Prevention, and Policy, 2:*32. doi:10.1186/1747-597X-2-32

O'Donnell, J. P., McCann, K. K., & Pluth, S. (2001). Assessing adult ADHD using a self-report symptom checklist. *Psychological Reports, 88,* 871–881. doi:10.2466/PR0.88.3.871-881

Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition, *J Exp Psychology. General*, *134,* 368–387.

Oken, B. S., Kishiyama, S. S., & Salinsky, M. C. (1995). Pharmacologically induced changes in arousal: effects on behavioral and electrophysiologic measures of alertness and attention. *Electroencephalography and Clinical Neurophysiology*, *95*(5), 359–371.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, *28*(10), 1021–1034.

Partridge, B., Bell, S., Lucke, J., and Hall, W. (2013). Australian university students' attitudes towards the use of prescription stimulants as cognitive enhancers: perceived patterns of use, efficacy and safety. *Drug Alcohol Rev*. 32, 295–302. doi: 10.1111/dar.12005

Pauls, A. M., O'Daly, O. G., Rubia, K., Riedel, W. J., Williams, S. C. R., & Mehta, M. A. (2012). Methylphenidate Effects on Prefrontal Functioning During Attentional-Capture and Response Inhibition. *Biological Psychiatry*, *72*(2), 142–149.

Paulson, P.E., Robinson, T.E. (1995). Amphetamine-Induced Time-Dependent Sensitization of Dopamine Neurotransmission in the Dorsal and Ventral Striatum: A Microdialysis Study in Behaving Rats. *Synapse*, *19*, 56–65.

Peterkin, A. L., Crone, C. C., Sheridan, M. J., Wise, T. N., Peterkin, A. L., Crone, C. C., Sheridan, M. J., et al. (2011). Cognitive Performance Enhancement : Misuse or Self-Treatment ? doi:10.1177/1087054710365980

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine, 26,* 4544-4562.

Peters, J. L., Sutton, A. J., Jones, D. R., & Abrams, K. R., Rushton, L., & Moreno, S. G. (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Journal of the Royal Statistical Society A, 173*(3), 575-591.

Pintrich, P., R., Smith, S., Garcia, T. & McKeachie, W. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ).* Washington, DC – Office of Educational Research and Improvement.

Postle, B. R., Desposito, M., Corkin, S., (2005). Effects of verbal and nonverbal interference on spatial and object visual working memory. *Memory & cognition*, *33*, 203-12.

Postle, B. R., Jonides, J., Smith, E. E., Corkin, S., & Growdon, J. H. (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. *Neuropsychology*, *11*(2), 171–179.

Poulin, C. (2001). Medical and nonmedical stimulant use among adoles- cents: From sanctioned to unsanctioned use. *CMAJ: Canadian Medical Association Journal/Journal de l'Association Medicale Canadienne, 165,* 1039–1044.

Poulin, C. (2007). From attention-deficit/hyperactivity disorder to medical stimulant use to the diversion of prescribed stimulants to non-medical stimulant use: connecting the dots. *Addiction*, *102*(5), 740–751.

Rabiner, D. L., Anastopoulos, A. D., Costello, E. J., Hoyle, R. H., McCabe, S. E., & Swartzwelder, H. S. (2009). Motives and perceived consequences of nonmedical ADHD medication use by college students: are students treating themselves for attention problems? *Journal of attention disorders*, 13(3), 259–70. doi:10.1177/1087054708320399

Rabiner, D. L., Anastopoulos, A. D., Costello, E. J., Hoyle, R. H., Swartzwelder, H. S., Costello, E. J., Hoyle, R. H., et al. (2010). Predictors of Nonmedical ADHD Medication Use by College Students. doi:10.1177/1087054709334505

Ramasubbu, R., Singh, H., Zhu, H., & Dunn, J. F. (2012). Methylphenidate-mediated reduction in prefrontal hemodynamic responses to working memory task: a functional near-infrared spectroscopy study. *Human Psychopharmacology: Clinical and Experimental*, *27*(6), 615–621.

Rasmussen, N. (2008). *On Speed: The Many Lives of Amphetamine*. New York, NY: New York University Press.

Raven, J. C., (1976). *Standard progressive matrices: Sets A, B, C, D & E.* Oxford: Oxford Psychologists Press.

Repantis, D., Schlattmann, P., Laisney, O., & Heuser, I. (2010). Modafinil and methylphenidate for neuroenhancement in healthy individuals: A systematic review. *Pharmacological Research*, *62*(3), 187–206.

Roache, R. (2009). Enhancement and Cheating. *Expositions*, 2(2), 153–156. doi:10.1558/expo.v2i2.153

Robbins, T. W., Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: Monoaminergic modulation. *Annu Rev Neurosci, 32,* 267–287.

Roediger, H. L., & Pyc, M. A. (2012). Journal of Applied Research in Memory and Cognition Inexpensive techniques to improve education : Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, 1(4), 242–248. doi:10.1016/j.jarmac.2012.09.002

Rosenthal, R., & Rubin, D. B. (2003). R-equivalent: A Simple Effect Size Indicator. *Psychological Methods*, *8*(4), 492–496.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. John Wiley & Sons.

Sahakian, B. J., & Morein-Zamir, S. (2011). Neuroethical issues in cognitive enhancement. *Journal of Psychopharmacology*, *25*(2), 197–204.

Sahakian, B., and Morein-Zamir, S. (2007). Professor's little helper. *Nature* 450, 1157–1159. doi: 10.1038/4501157a

Schatz, A. M. (2001). Sensitivity and specificity of a computerized test of attention in the diagnosis of Attention-Deficit/Hyperactivity Disorder. *Assessment,* 8(4), 357 -365

Schenk, S., Pattridge, B. (1997). Sensitization and tolerance in psychostimulant self-administration. Pharmacology, *Biochemistry and Behavior*. 57, 543-550.

Schmeck, R. R. , Geisler-Brenstein, E. & Cercy, S.P. (1991). Self-concept and Learning: The revised Inventory of Learning Processes. *Educational Psychology,* 11 (*3-4*), 343-362

Schmedtje, J. F., Jr, Oman, C. M., Letz, R., & Baker, E. L. (1988). Effects of scopolamine and dextroamphetamine on human performance. *Aviation, Space, and Environmental Medicine*, *59*(5), 407–410.

Servan-Schreiber, D., Carter, C. S., Bruno, R. M., & Cohen, J. D. (1998). Dopamine and the mechanisms of cognition: Part II. D-amphetamine effects in human subjects performing a selective attention task. *Biological Psychiatry*, *43*(10), 723–729.

Silber, B. Y., Croft, R. J., Papafotiou, K., & Stough, C. (2006). The acute effects of d-amphetamine and methamphetamine on attention and psychomotor performance. *Psychopharmacology*, *187*(2), 154–169.

Simmons, J.P., Nelson, L.D. & Simonsohn, U. (2011). False-positive psychology. *Psychological Science,* 22 (11).

Sitzmann, T., Ely, K., & Group, F. M. (2011). A Meta-Analysis of Self-Regulated Learning in Work-Related Training and Educational Attainment : What We Know and Where We Need to Go, 137(3), 421–442. doi:10.1037/a0022777

Smith, G. M.; Beecher, Henry K. (1964). Drugs and judgment: Effect of amphetamine and secobarbital on self-evaluation. *J Psychology*: Interdisciplinary and Applied, 58, 397-405.

Smith, M. E., Farah, M. J. (2011). Are prescription stimulants “smart pills”? The epidemiology and cognitive neuroscience of prescription stimulant use by normal healthy individuals. *Psychological Bulletin*, *137*, 717–741.

Sockol, L. E., Epperson, C. N., & Barber, J. P. (2011). A meta-analysis of treatments for perinatal depression. *Clinical Psychology Review*, *31*(5), 839–849.

Soetens, E., Hueting, J. E., Casaer, S., & D’Hooge, R. (1995). Effect of amphetamine on long-term retention of verbal material. *Psychopharmacology*, *119*(2), 155–162.

Sofuoglu, M., Waters, A. J., Mooney, M., & Kosten, T. (2008). Riluzole and d-amphetamine interactions in humans. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *32*(1), 16–22.

Son, L. K., & Kornell, N. (2009). Simultaneous decisions at study: time allocation, ordering, and spacing. *Metacognition and Learning*, 4(3), 237–248. doi:10.1007/s11409-009-9049-1

Steer, R. A, Ball, R., Ranieri, W. F., & Beck, A. T. (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. *Journal of clinical psychology*, 55(1), 117–28.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654.

Storch, E. a, Roberti, J. W., & Roth, D. a. (2004). Factor structure, concurrent validity, and internal consistency of the Beck Depression Inventory-Second Edition in a

sample of college students. *Depression and anxiety*, 19(3), 187–9. doi:10.1002/da.20002

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.

Studer, P., Wangler, S., Diruf, M. S., Kratz, O., Moll, G. H., & Heinrich, H. (2010). ERP effects of methylphenidate and working memory load in healthy adults during a serial visual working memory task. *Neuroscience Letters*, *482*(2), 172–176.

Tablot, M. (2009). Brain gain. The New Yorker. Retrieved from: http://www.iricss.org/fa/NewsRelease/News/Documents/Brain_Gain_The_New_Yorker.pdf

Teter, C. J., Falone, A. E., Cranford, J.A., Boyd, C. J., and McCabe, S. E. (2010). Nonmedical use of prescription stimulants and depressed mood among college students: frequency and routes of administration. *J. Subst. Abuse Treat.* 38, 292–298. doi: 10.1016/j.jsat.2010.01.005

Teter, C. J., McCabe, S. E., Cranford, J. A., Boyd, C. J., and Guthrie, S. K. (2005). Prevalence and motives for illicit use of prescription stimulants in an undergraduate student sample. *J. Am. College Health* 53, 253–262. doi: 10.3200/JACH.53.6.253-262

The Psychological Corporation., 2002. Updated WAISIII-WMS-III technical manual. San Antonio, TX.Tablot, M., 2009. Brain gain. The New Yorker.

Theunissen, E. L., Elvira, J. de la A., van den Bergh, D., & Ramaekers, J. G. (2009). Comparing the stimulant effects of the H1-antagonist fexofenadine with 2

psychostimulants, modafinil and methylphenidate. *Journal of Clinical Psychopharmacology*, *29*(5), 439–443.

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & David, H. (2009). Worth the " EEfRT "? The Effort Expenditure for Rewards Task as an Objective Measure of Motivation and Anhedonia, 4(8), 1–9. doi:10.1371/journal.pone.0006598

Unrug, A., Coenen, A., & van Luijtelaar, G. (1997). Effects of the Tranquillizer Diazepam and the Stimulant Methylphenidate on Alertness and Memory. *Neuropsychobiology*, *36*(1), 42–48.

Verdejo-García, A., Lawrence, A. J., & Clark, L. (2008). Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. *Neuroscience and biobehavioral reviews*, 32(4), 777–810. doi:10.1016/j.neubiorev.2007.11.003

Volkow, N. D., Wang, G. J., Fowler, J. S., Telanfg F., Maynard, L., Logan, J., et al. (2004). Evidence that methylphenidate enhances the saliency of a mathematical task by increasing dopamine in the human brain. *Am. J. Psychiatry* 161, 1173–1180. doi: 10.1176/appi.ajp.161.7.1173

Volkow, N. D., Wang, G.-J., Fowler, J. S., Gatley, S. J., Logan, J., Ding, Y.-S., … Pappas, N. (1998). Dopamine transporter occupancies in the human brain induced by therapeutic doses of oral methylphenidate. *American Journal of Psychiatry*, *155*(10), 1325–1331.

Volkow, N. D., Wang, G.-J., Kollins, S. H., Wigal, T. L., Newcorn, J. H., Telang, F., … Ma, Y. (2009). Evaluating dopamine reward pathway in ADHD: Clinical implications. *JAMA*, *302*(10), 1084–1091.

Vrecko, S. (2013). Just how cognitive is "cognitive enhancement"? On the significance of emotions in university students' experiences with study drugs. *AJOB Neuroscience*, *4*(1), 4–12.

Vree, T. B., & Van Rossum, J. M. (1970). Kinetics of metabolism and excretion of amphetamines in man. *Amphetamines and related compounds*, 165-190.

Wardle, M. C., and de Wit, H. (2012). Effects of amphetamine on reactivity to emotional stim- uli. *Psychopharmacology* 220, 143–153. doi: 10.1007/s00213-011-2498-7

Wardle, M. C., Hart, A. B., Palmer, A. A., & de Wit, H. (2013). Does COMT genotype influence the effects of d-amphetamine on executive functioning? *Genes, Brain and Behavior*, *12*(1), 13–20.

Wardle, M. C., Treadway, M. T., Mayo, L. M., Zald, D. H., and de Wit, H. (2011). Amping up effort: effects of d-Amphetamine on human effort-based decision-making. *J. Neurosci.* 31, 16597–16602. doi: 10.1523/JNEUROSCI.4387-11.2011

Weyandt, L. L., Janusis, G., Wilson, K. G., Verdi, G., Paquin, G., Lopes, J., et al. (2009). Nonmedical prescription stimulant use among a sample of college students: relation- ship with psychological variables. *J. Atten. Disord.* 3, 284–296. doi: 10.1177/1087054709342212

Wilens, T. E., Adler, L. A., Adams, J., Sgambati, S., Rotrosen, J., Sawtelle, R., … Fusillo, S. (2008). Misuse and diversion of stimulants prescribed for ADHD: A systematic review of the literature. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(1), 21–31.

Willett, R. A. (1962). The effect of a stimulant and a depressant drug on the serial rote learning of nonsense syllables. *Psychopharmacologia*, *3*(1), 23–34.

Zeeuws, I., & Soetens, E. (2007). Verbal memory performance improved via an acute administration of D-amphetamine. *Human Psychopharmacology: Clinical and Experimental*, *22*(5), 279–287.

Zeeuws, I., Deroost, N., & Soetens, E. (2010a). Effect of an acute d-amphetamine administration on context information memory in healthy volunteers: evidence from a source memory task. *Human Psychopharmacology: Clinical and Experimental*, *25*(4), 326–334.

Zeeuws, I., Deroost, N., & Soetens, E. (2010b). Verbal memory improved by D-amphetamine: influence of the testing effect. *Human Psychopharmacology: Clinical and Experimental*, *25*(5), 377–387.